

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:31:33 ; Search time 69.2093 Seconds
(without alignments)
111.765 Million cell updates/sec

Title: US-10-612-090-14

Perfect score: 104
Sequence: 1 SSVLDGYSPPNREPLTGNS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	20	8	ADJ66741 Human Muc
2	104	100.0	95	8	ADP81172 Protein o
3	104	100.0	108	8	ADJ66729 Human Muc
4	104	100.0	108	8	ADJ66740 Human Muc
5	104	100.0	109	8	ADP81173 Protein o
6	104	100.0	174	7	ADM30756 Human can
7	104	100.0	284	6	ABU54859 Human CA1
8	104	100.0	438	5	AAE12623 Human gen
9	104	100.0	438	5	ABP30973 Partial p
10	104	100.0	438	7	ADA08636 Human O77
11	104	100.0	438	7	ADF08979 Secreted
12	104	100.0	439	6	ABU54709 Human CA1
13	104	100.0	526	4	AAE12634 Human gen
14	104	100.0	748	8	ADJ94302 CA 125/O7
15	104	100.0	809	8	ADJ94303 CA 125/O7
16	104	100.0	1148	4	AAE5836 Human pro
17	104	100.0	1148	4	ABBS0283 HOST-1 ov
18	104	100.0	1148	5	ABG36380 Human ova
19	104	100.0	1148	5	ABP30964 Truncated
20	104	100.0	1148	5	ABP30969 Clone FLJ
21	104	100.0	1148	6	ABU54692 Human CA1
22	104	100.0	1148	6	ABU00145 Human can
23	104	100.0	1148	7	ADA08632 Human O77
24	104	100.0	1148	7	ADA08611 Human O77
25	104	100.0	1148	7	ADF08954 Secreted

26	104	100.0	1148	7	ADF08975	Adf08975 Secreted
27	104	100.0	1148	7	ADG46241	Adg46241 Human ova
28	104	100.0	1148	7	ADM30754	Adm30754 Human can
29	104	100.0	1148	8	ADH41326	Adh41326 Human ova
30	104	100.0	1148	8	ADQ29705	Adq29705 Human col
31	104	100.0	1156	5	ABP30965	Abp30965 Protein w
32	104	100.0	1156	7	ADA08612	Ada08612 Human O77
33	104	100.0	1156	7	ADF08955	Adf08955 Secreted
34	104	100.0	1156	7	ADG46242	Adg46242 Human ova
35	104	100.0	1307	8	ADP10190	Adp10190 Human pro
36	104	100.0	1889	6	AAE34700	Aae34700 Protein e
37	104	100.0	1890	5	ABG36381	Abg36381 Human ova
38	104	100.0	1890	8	ADM12352	Adm12352 Human CA1
39	104	100.0	2234	8	ADJ66747	Adj66747 Human Muc
40	104	100.0	2248	6	ABU54693	Abu54693 Human CA1
41	104	100.0	11721	6	ABU54721	Abu54721 Human CA1
42	104	100.0	22152	8	ADR72871	Adr72871 Human ova
43	104	100.0	22157	8	ADP84155	Adp84155 Human CA1
44	101	97.1	1890	8	ADO38574	Ado38574 Ovarian c
45	100	96.2	318	5	ABP31025	Abp31025 Amino aci

ALIGNMENTS

RESULT 1

ADJ66741 standard; peptide; 20 AA.

ADJ66741; 06-MAY-2004 (first entry)

Human Muc1/Muc16 protein-related peptide SegID14.

monoclonal antibody; epitope; non-shed extracellular portion; shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy; breast cancer; ovarian cancer.

OS Homo sapiens.

XX WO2004005470-A2.

XX PD 15-JAN-2004.

XX PF 03-JUL-2003; 2003WO-US020907.

XX PR 03-JUL-2002; 2002US-0393094P.

XX PA (IMMU-) IMMUNOGEN INC.

XX PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K; Water CA; WPI; 2004-091350/09.

XX DR MPI; 2004-091350/09.

XX PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing, treating or monitoring malignancies, such as breast or ovarian cancer.

XX PS Claim 14; SEQ ID NO 14; 113pp; English.

XX CC This invention relates to a novel isolated monoclonal antibody that specifically binds to an epitope of a non-shed extracellular portion of a shed antigen or of human Muc1 or Muc16 protein. The invention may be useful for the development of compounds with a cytostatic or for gene therapy. The composition and methods are useful in diagnosing, treating or monitoring malignancies, such as breast or ovarian cancer. The present CC sequence is that of a peptide which was used for raising antibodies to the extracellular, non-shed region of Muc16 and which is claimed in the CC specification.

XX CC Sequence 20 AA;

XX SQ

Query Match 100.0%; Score 104; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2,7e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SSVLVGYSPPNRNEPLTGS 20
 |||
 1 SSVLVGYSPPNRNEPLTGS 20

Db 1 SSVLVGYSPPNRNEPLTGS 20

RESULT 2
 ADP81172
 ID ADP81172 standard; protein; 95 AA.

AC ADP81172;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Protein of human ovarian specific gene, SEQ ID NO 206.
 XX
 KM normal; neoplastic; ovarian; ovarian specific nucleic acid; OSNA;
 KM metastatic; cancer; vaccine; cytostatic; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004053079-A2.
 XX
 PD 24-JUN-2004.
 XX
 PF 08-DEC-2003; 2003WO-US038855.
 XX
 PR 06-DEC-2002; 2002US-0431301P.
 PR 06-DEC-2002; 2002US-0431321P.
 PR 30-JUN-2003; 2003US-0484584P.
 PR 07-NOV-2003; 2003US-0518607P.
 XX
 PA (D1AD-) DIADEXUS INC.
 XX
 PI Macina RA, Turner LR, Sun Y, Liu S, Chen H;
 PI
 DR WPI; 2004-468850/44.
 DR N-PSDB; ADP81027.
 XX
 PT New ovarian specific nucleic acid molecules and polypeptides useful for
 PT diagnosing, preventing or treating ovarian cancer, for producing
 PT transgenic animals or cells, or for research purposes.
 XX
 PS Claim 12; SEQ ID NO 206; 754pp; English.

The invention relates to novel isolated nucleic acid molecules and polypeptides present in normal and neoplastic ovarian cells. These comprise a nucleic acid sequence encoding any of the 167 amino acid sequences (e.g. 438, 237 or 233 amino acids) fully defined in the specification (SEQ. ID NOS: ADP81095 to ADP81261) and comprises any of the 128 nucleotide sequences (e.g. 4798, 1494 or 1691 bp) fully defined in the specification (SEQ. ID NOS: ADP80967 to ADP81094). The invention further comprises: a method for determining the presence of a ovarian specific nucleic acid (OSNA) in a sample; a vector comprising the above nucleic acid molecule; a host cell comprising the vector; a method for producing a polypeptide encoded by the above nucleic acid molecule; a polypeptide encoded by the nucleic acid molecule cited above; an antibody or its fragment that specifically binds to the above polypeptide; a method for determining the presence of an ovarian specific protein in a sample; a method for diagnosing or monitoring the presence and metastases of ovarian cancer in a patient; a kit for detecting a risk of cancer or presence of cancer in a patient; the kit comprising a means for determining the presence of the above nucleic acid molecule or polypeptide; a method of treating a patient with ovarian cancer; and a vaccine comprising the above polypeptide or nucleic acid encoding the polypeptide. The isolated nucleic acid molecules and polypeptides have cytostatic activity. The isolated polypeptides may be used to create a vaccine. The isolated nucleic acid molecules and polypeptides can be used for diagnosing or monitoring the presence and metastases of ovarian cancer and treating ovarian cancer. This sequence represents the protein

CC of an ovarian specific gene of the invention.
 XX
 SQ Sequence 95 AA;

Query Match 100.0%; Score 104; DB 8; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1,6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SSVLVGYSPPNRNEPLTGS 20
 |||
 17 SSVLVGYSPPNRNEPLTGS 36

Db 17 SSVLVGYSPPNRNEPLTGS 36

RESULT 3
 ADJ66729
 ID ADJ66729 standard; protein; 108 AA.

AC ADJ66729;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human Muc16 epitope amino acid sequence.
 XX
 KM monoclonal antibody; epitope; non-shed extracellular portion;
 KM shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KM breast cancer; ovarian cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO2004005470-A2.
 XX
 PD 15-JAN-2004.
 XX
 PF 03-JUL-2003; 2003WO-US020907.
 XX
 PR 03-JUL-2002; 2002US-0393094P.
 XX
 PA (IMMU-) IMMUNOGEN INC.
 XX
 PI Payne G, Chittenden T, Goldmakher V, Chun P, Snelder-Mulready K,
 PI Vater CA;
 PI
 DR WPI; 2004-091350/09.
 XX
 PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
 PT treating or monitoring malignancies, such as breast or ovarian cancer.
 XX
 PS Claim 13; SEQ ID NO 2; 113pp; English.

This invention relates to a novel isolated monoclonal antibody that specifically binds to an epitope of a non-shed extracellular portion of a shed antigen or of human Muc1 or Muc16 protein. The invention may be useful for the development of compounds with a cytostatic or for gene therapy. The composition and methods are useful in diagnosing, treating or monitoring malignancies, such as breast or ovarian cancer. The present sequence is that of a Muc epitope of the invention.

Sequence 108 AA;

Query Match 100.0%; Score 104; DB 8; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1,7e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SSVLVGYSPPNRNEPLTGS 20
 |||
 86 SSVLVGYSPPNRNEPLTGS 105

Db 86 SSVLVGYSPPNRNEPLTGS 105

RESULT 4
 ADJ66740
 ID ADJ66740 standard; protein; 108 AA.

AC ADJ66740;

XX 06-MAY-2004 (first entry)
XX
XX
DE Human Muc16 GST fusion protein amino acid sequence SeqID13.
XX monoclonal antibody; epitope; non-shed extracellular portion;
XX shed antigen; human; Muc1; Muc16; cytosolic; gene therapy; malignancy;
XX breast cancer; ovarian cancer.
XX Homo sapiens.
XX WO2004005470-A2.
XX 15-JAN-2004.
XX
XX 03-JUL-2003; 2003WO-US020907.
XX
XX 03-JUL-2002; 2002US-0393094P.
XX (IMMU-) IMMUNOGEN INC.
XX
XX Payne G, Chittenden T, Goldmacher V, Chun P, Sneider-Mulready K;
XX Vater CA;
XX MPI; 2004-091350/09.
XX
XX New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
XX treating or monitoring malignancies, such as breast or ovarian cancer.
XX
XX Disclosure; SEQ ID NO 13; 113pp; English.
XX
XX This invention relates to a novel isolated monoclonal antibody that
XX specifically binds to an epitope of a non-shed extracellular portion of a
XX shed antigen or of human Muc1 or Muc16 protein. The invention may be
XX useful for the development of compounds with a cytostatic or for gene
XX therapy. The composition and methods are useful in diagnosing, treating
XX or monitoring malignancies, such as breast or ovarian cancer. The present
XX sequence is that of the human Muc16 GST fusion protein which was used in
XX the exemplification of the invention.
XX
XX Sequence 108 AA;
SQ
Query Match 100.0%; Score 104; DB 8; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVLDGYSPPNREPLTGNS 20
DB 86 SSVLDGYSPPNREPLTGNS 105
RESULT 5
ADP81173
ID ADP81173 standard; protein; 109 AA.
XX
XX ADP81173;
XX
XX 09-SEP-2004 (first entry)
XX
XX Protein of human ovarian specific gene, SEQ ID NO 207.
XX
XX normal; neoplastic; ovarian; ovarian specific nucleic acid; OSNA;
XX metastatic; cancer; vaccine; cytosolic; human.
XX
XX Homo sapiens.
XX
XX WO2004053079-A2.
XX
XX 24-JUN-2004.
XX
XX 08-DEC-2003; 2003WO-US038855.
XX
XX 06-DEC-2002; 2002US-0431301P.
XX
XX

PR 06-DEC-2002; 2002US-0431321P.
PR 30-JUN-2003; 2003US-0484584P.
PR 07-NOV-2003; 2003US-0518607P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Macina RA, Turner LR, Sun Y, Liu S, Chen H;
XX N-PSDB; ADP81027.
XX
XX MPI; 2004-468550/44.
XX
XX New ovarian specific nucleic acid molecules and polypeptides useful for
XX diagnosing, preventing or treating ovarian cancer, for producing
XX transgenic animals or cells, or for research purposes.
XX
XX Claim 12; SEQ ID NO 207; 754pp; English.
XX
XX The invention relates to novel isolated nucleic acid molecules and
XX polypeptides present in normal and neoplastic ovarian cells. These
XX comprise a nucleic acid sequence encoding any of the 167 amino acid
XX sequences (e.g. 438, 237 or 233 amino acids) fully defined in the
XX specification (SEQ. ID NOS: ADP81095 to ADP81261) and comprises any of
XX the 128 nucleotide sequences (e.g. 4798, 1494 or 1691 bp) fully defined
XX in the specification (SEQ. ID NOS: ADP80967 to ADP81094). The invention
XX further comprises: a method for determining the presence of an ovarian
XX specific nucleic acid (OSNA) in a sample; a vector comprising the above
XX nucleic acid molecule; a host cell comprising the vector; a method for
XX producing a polypeptide encoded by the above nucleic acid molecule; a
XX polypeptide encoded by the nucleic acid molecule cited above; an antibody
XX or its fragment that specifically binds to the above polypeptide; a
XX method for determining the presence of an ovarian specific protein in a
XX sample; a method for diagnosing or monitoring the presence and metastases
XX of ovarian cancer in a patient; a kit for detecting a risk of cancer or
XX presence of cancer in a patient; the kit comprising a means for
XX determining the presence of the above nucleic acid molecule or
XX polypeptide; a method of treating a patient with ovarian cancer; and a
XX vaccine comprising the above polypeptide or nucleic acid encoding the
XX polypeptide. The isolated nucleic acid molecules and polypeptides have
XX cytostatic activity. The isolated polypeptides may be used to create a
XX vaccine. The isolated nucleic acid molecules and polypeptides can be used
XX for diagnosing or monitoring the presence and metastases of ovarian
XX cancer and treating ovarian cancer. This sequence represents the protein
XX of an ovarian specific gene of the invention.
XX
XX Sequence 109 AA;
SQ
Query Match 100.0%; Score 104; DB 8; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVLDGYSPPNREPLTGNS 20
DB 31 SSVLDGYSPPNREPLTGNS 50
RESULT 6
ADM30756
ID ADM30756 standard; protein; 174 AA.
XX
XX ADM30756;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human cancer linked protein referred to as 192292 SeqID 4.
XX
XX human; cancer-related gene; screening assay; immuno-conjugate;
XX cytostatic; immunostimulant; cancer; ovarian cancer.
XX
XX Homo sapiens.
XX
XX WO2003075854-A2.
XX
XX 18-SEP-2003.
XX
XX

XX WP1: 2003-093013/08.

XX New CA125 molecules, useful as a gold standard for detecting and

PT monitoring the presence of CA125 antigen which can be used for

PT diagnosing, monitoring or treating patients with cancer or for developing

PT vaccine against cancer.

XX

XX Claim 1: Fig 9b, 694pp; English.

XX

CC The invention relates to a CA125 protein comprising: (a) an extracellular

CC amino terminal domain; (b) an amino terminal extension; (c) a multiple

CC repeat domain; and (d) a carboxy terminal domain. The extracellular amino

CC terminal domain comprises 5 genomic exons, the amino terminal extension

CC comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and

CC the carboxy terminal domain comprises a transmembrane anchor with a short

CC cytoplasmic domain, and further comprises 9 genomic exons. The gene for

CC CA125 is located on human chromosome 19q 13.2. Also included are isolated

CC CA125 repeat domains, nucleic acids (including variants, homologues and

CC degenerate versions) encoding CA125 proteins or repeat units, a vector

CC comprising the nucleic acid, a cultured cell comprising the vector, a

CC method of expressing CA125 antigen in a cell, the amino acid sequences of

CC the CA125 repeat units (or their variants, fragments or sequences 50%

CC identical to them), a purified antibody that selectively binds to an

CC epitope in the receptor-binding domain of CA125 antigen (comprising

CC detecting and monitoring the presence of CA125 antigen (comprising

CC recombinant CA125 having at least one repeat unit of the CA125 repeat

CC domain including epitope binding sites), a therapeutic vaccine to treat

CC mammals with elevated CA125 antigen levels or at risk of developing a

CC disease or disease recurrence associated with elevated CA125 antigen

CC levels (comprising recombinant CA125 repeat domains including epitope

CC binding sites) and an antisense oligonucleotide that inhibits the

CC expression of CA125. The CA125 molecule, particularly the multiple repeat

CC domains are useful as a gold standard for detecting and monitoring the

CC presence of CA125 antigen, which can be used for diagnosing, monitoring

CC or treating patients with ovarian cancer and other carcinomas where CA125

CC is expressed. The molecules are also useful for developing a vaccine

CC against cancer. The present sequence is a CA125 repeat protein

CC

XX

XX Sequence 284 AA:

XX

XX

XX Query Match 100.0%; Score 104; DB 6; Length 284;

XX Best Local Similarity 100.0%; Pred. No. 5.6e-08;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

XX

OY 1 SSVLVDSYSPNRNPEPLTGS 20

DB 206 SSVLVDSYSPNRNPEPLTGS 225

XX

RESULT 8

XX AAE12623

XX ID AAE12623 standard; protein; 438 AA.

XX

XX AAE12623;

XX

XX 03-JAN-2002 (first entry)

XX

XX Human gene 1 encoded secreted protein H1PA230.

XX

XX Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;

XX respiratory system disorder; asthma; haematopoietic disorder; skin aging;

XX immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;

XX rheumatoid arthritis; inflammation; neurological disorder; septic shock;

XX Alzheimer's disease; Parkinson's disease; diabetes; idiopathic disorder;

XX atherosclerosis; cardiovascular disorder; gastrointestinal disorder;

XX epithelial cell proliferation; transplantation; chemotaxis; infection;

XX food additive; wound healing; endocrine disorder; kidney disorder;

XX gene therapy; cytoskeletal.

XX

XX Homo sapiens.

XX

XX WO200170804-A1.

XX 27-SEP-2001.
PD
XX
PF 16-MAR-2001; 2001WO-US008585.
XX
PR 17-MAR-2000; 2000US-0190076P.
XX 23-AUG-2000; 2000US-0227009P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Birse CE, Rosen CA;
XX WPI; 2001-639119/73.
XX DR N-PSDB; AAD20618.
XX
PT Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
PT metastases.
XX
XX
XX Claim 11; Page 412-414; 427pp; English.
XX
XX AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
CC ovarian cancer associated protein (collectively known as ovarian cancer
CC antigen) genes, and AAE12623-AAE12629 represent the proteins they
CC encode. AAE12630-AAE12638 represent human ovarian cancer antigenic
CC fragments or variants. Ovarian cancer antigens and their corresponding
CC DNAs are used in the prevention, diagnosis and treatment of diseases
CC associated with their inappropriate expression. These disorders include
CC proliferative disorders, cancer, tumours, respiratory system disorders,
CC asthma, haematopoietic disorders, diseases of the immune system, AIDS,
CC skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
CC inflammation, allergies, neurological disorders (e.g., Alzheimer's
CC disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, endocrine disorders and infections. The
CC protein can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, to identify their cognate ligands or binding
CC partners, in chemotaxis and can be used as a food additive. Antibodies
CC specific for a protein of the invention can be used in alleviating
CC symptoms associated with the disorders mentioned above and in diagnostic
CC immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
CC the invention is used in gene therapy. The present sequence represents a
CC human ovarian cancer antigen of the invention
XX
XX
SQ Sequence 438 AA;
Query Match 100.0%; Score 104; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 9.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVLVGYSPPNNEPLTGN 20
DB 360 SSVLVGYSPPNNEPLTGN 379
RESULT 9
ABP30973
ID ABP30973 standard; protein; 438 AA.
XX
XX ABP30973;
XX
DT 02-JUL-2002 (first entry)
XX
XX Partial protein sequence of clone 0772P partial cDNA sequence 4.
DE
XX
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
KW
XX Homo sapiens.
OS
XX WO200206317-A2.
PN
XX
XX 24-JAN-2002.
PD

XX
XX 17-JUL-2001; 2001WO-US022635.
PF
XX
XX 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
XX 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
XX 18-JUN-2001; 2001US-00844441.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone B;
XX
XX WPI; 2002-164781/21.
XX DR N-PSDB; ABN72979.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
XX Example 13; Page 357; 408pp; English.
XX
XX This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX
XX
SQ Sequence 438 AA;
Query Match 100.0%; Score 104; DB 5; Length 438;
Best Local Similarity 100.0%; Pred. No. 9.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVLVGYSPPNNEPLTGN 20
DB 360 SSVLVGYSPPNNEPLTGN 379
RESULT 10
ADA08636
ID ADA08636 standard; protein; 438 AA.
XX
XX ADA08636;
XX
DT 06-NOV-2003 (first entry)
XX
XX Human 0772P partial protein #6.
DE
XX
XX human; gene therapy; ovarian cancer; cancer.
KW
XX Homo sapiens.
OS
XX US2003091580-A1.
PN
XX
XX 15-MAY-2003.
PD
XX
XX 17-JUL-2001; 2001US-00907969.
PF
XX
XX 18-JUN-2001; 2001US-00844441.
PR
XX
XX (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.

PA (HILL/) HILL P.
 PA (ALBO/) ALBONE E.
 XX
 PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedrick TS, Carter D, Hill P, Albone E;
 XX
 DR WPI: 2003-532352/50.
 XX
 PT New isolated 0772p polypeptides and polynucleotides, useful in gene
 PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
 PT cancer.
 XX
 PS Example 13; SEQ ID NO 483; 371pp; English.
 XX
 CC The invention relates to an isolated 0772p polypeptide, which has the
 CC structure fully defined in the specification. The composition containing
 CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
 CC or antigen presenting cells are useful for stimulating an immune response
 CC and treating ovarian cancer. Detecting the presence of the
 CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
 CC carcinoma cDNAs and protein cDNAs were identified using microarray
 CC technology. The present sequence represents a human ovarian carcinoma
 CC antigen.
 CC
 SQ Sequence 438 AA;

Query Match 100.0%; Score 104; DB 7; Length 438;
 Best Local Similarity 100.0%; Pred. No. 9.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLVGYSPPNNEPLTGNS 20
 |||||
 Db 360 SSVLVGYSPPNNEPLTGNS 379

RESULT 11

ADP08979
 ID ADF08979 standard; protein; 438 AA.

AC ADF08979;
 XX

DT 12-FEB-2004 (first entry)
 XX

DE Secreted ovarian carcinoma antigen seqid 483.
 XX

XX gene therapy; protein therapy; vaccine; antibody inhibition;
 XX breast cancer; restorative therapy; diagnostic agent; immunoassay;
 KW secreted ovarian carcinoma antigen.

OS Homo sapiens.
 XX

PN US2003124140-A1.
 XX

PD 03-JUL-2003.
 XX

PF 17-JUL-2002; 2002US-00198053.
 XX

PR 17-DEC-1998; 98US-00215681.
 XX

PR 17-DEC-1998; 98US-00216003.
 XX

PR 23-JUN-1999; 99US-00338933.
 XX

PR 24-SEP-1999; 99US-00404879.
 XX

PR 17-JUL-2000; 2000US-00617747.
 XX

PR 10-AUG-2000; 2000US-00636801.
 XX

PR 20-SEP-2000; 2000US-00667857.
 XX

PR 04-APR-2001; 2001US-00827271.
 XX

PR 18-JUN-2001; 2001US-00884441.
 XX

PR 17-JUL-2001; 2001US-00907969.
 XX

PA (CORI-) CORIXA CORP.
 XX

PI Bangur CS, Retter MW, Fanger GR, Hill P;
 XX

PT WPI: 2003-897152/82.
 XX

XX
 PT Oncogenic nucleic acids useful for the prevention, diagnosis and
 PT treatment of breast cancer.
 XX
 PS Example 13; SEQ ID NO 483; 399pp; English.
 XX

CC The invention describes nucleic acids (I) and the polypeptides (II) they
 CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
 CC treating diseases related to their aberrant expression i.e. breast
 CC cancers. For example, (I) and (II) may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of (II) by expressing
 CC inactive proteins or to supplement the patient's own production of (II).
 CC Additionally, (II) may be used to produce (II)', by inserting (I) into a
 CC host cell and culturing the cell to express the protein (II)'. (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. The host cell may also be used as antigens in the production of
 CC antibodies against (II) and in assays to identify modulators of (II)'s
 CC expression and activity. The anti-(II) antibodies, agonists and
 CC antagonists may be used to regulate expression and activity and as
 CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
 CC immunoassay). This sequence represents a secreted ovarian carcinoma
 CC antigen.
 CC
 SQ Sequence 438 AA;

Query Match 100.0%; Score 104; DB 7; Length 438;
 Best Local Similarity 100.0%; Pred. No. 9.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLVGYSPPNNEPLTGNS 20
 |||||
 Db 360 SSVLVGYSPPNNEPLTGNS 379

RESULT 12

ABU54709
 ID ABU54709 standard; protein; 439 AA.

AC ABU54709;
 XX

DT 12-MAR-2003 (first entry)
 XX

DE Human CA125 protein carboxy terminus.
 XX

XX Human; CA125; protein repeat; chromosome 19q 13.2; amino terminal domain;
 KW amino terminal extension; carboxy terminal domain; vaccine; cancer;
 KW ovarian cancer; carcinoma.
 XX

OS Homo sapiens.
 XX

PN WO200283866-A2.
 XX

PD 24-OCT-2002.
 XX

PF 12-APR-2002; 2002WO-US011734.
 XX

PR 17-APR-2001; 2001US-0284175P.
 XX

PR 19-JUN-2001; 2001US-0299380P.
 XX

PR 27-SEP-2001; 2001US-00965738.
 XX

PR 21-DEC-2001; 2001US-0345180P.
 XX

PA (UVAR-) UNIV ARKANSAS.
 XX

PI O'brien T, Beard J, Underwood L;
 XX

DR WPI: 2003-093013/08.
 XX

DR N-PSDB; ABX72616.
 XX

PT New CA125 molecules, useful as a gold standard for detecting and
 PT monitoring the presence of CA125 antigen which can be used for

PT diagnosing, monitoring or treating patients with cancer or for developing
PT vaccine against cancer.

PS Claim 23; Page 111; 694pp; English.

XX The invention relates to a CA125 protein comprising: (a) an extracellular
CC amino terminal domain; (b) an amino terminal extension; (c) a multiple
CC repeat domain; and (d) a carboxy terminal domain. The extracellular amino
CC terminal domain comprises 5 genomic exons, the amino terminal extension
CC comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and
CC the carboxy terminal domain comprises a transmembrane anchor with a short
CC cytoplasmic domain, and further comprises 9 genomic exons. The gene for
CC CA125 is located on human chromosome 19q 13.2. Also included are isolated
CC CA125 repeat domains, nucleic acids (including variants, homologues and
CC degenerate versions) encoding CA125 proteins or repeat units, a vector
CC comprising the nucleic acid, a cultured cell comprising the vector, a
CC method of expressing CA125 antigen in a cell, the amino acid sequences of
CC the CA125 repeat units (or their variants, fragments or sequences 50%
CC identical to them), a purified antibody that selectively binds to an
CC epitope in the receptor-binding domain of CA125 protein, a diagnostic for
CC detecting and monitoring the presence of CA125 antigen (comprising
CC recombinant CA125 having at least one repeat unit of the CA125 repeat
CC domain including epitope binding sites), a therapeutic vaccine to treat
CC mammals with elevated CA125 antigen levels or at risk of developing a
CC disease or disease recurrence associated with elevated CA125 antigen
CC levels (comprising recombinant CA125 repeat domains including epitope
CC binding sites) and an antisense oligonucleotide that inhibits the
CC expression of CA125. The CA125 molecule, particularly the multiple repeat
CC domains are useful as a gold standard for detecting and monitoring the
CC presence of CA125 antigen, which can be used for diagnosing, monitoring
CC or treating patients with ovarian cancer and other carcinomas where CA125
CC is expressed. The molecules are also useful for developing a vaccine
CC against cancer. The present sequence is a CA125 repeat protein

XX Sequence 439 AA;

Query Match 100.0%; Score 104; DB 6; Length 439;
Best Local Similarity 100.0%; Pred. No. 9, 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNRPITGNS 20
DB 361 SSVLDGYSPPNRPITGNS 380

RESULT 13

AAE12634
ID AAE12634 standard; protein; 526 AA.

AC AAE12634;

DT 03-JAN-2002 (first entry)

DE Human gene 1 encoded secreted protein fragment, SEQ ID NO: 30.

XX Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;
XX respiratory system disorder; asthma; haematopoietic disorder; skin aging;
XX immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
XX rheumatoid arthritis; inflammation; neurological disorder; septic shock;
XX Alzheimer's disease; Parkinson's disease; diabetes; angiogenic disorder;
XX atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
XX epithelial cell proliferation; transplantation; chemotaxis; infection;
XX food additive; wound healing; endocrine disorder; kidney disorder;
XX gene therapy; cytostatic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Mtec-difference 116

XX /label= Unknown

XX /note="xaa equals any of the naturally occurring L-amino
XX acids"

PN WO200170804-A1.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US008585.

XX 17-MAR-2000; 2000US-0190076P.

XX 23-AUG-2000; 2000US-0227009P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2001-639119/73.

PT Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
PT metastases.

PS Disclosure; Page 10; 427pp; English.

XX AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
CC ovarian cancer associated protein (collectively known as ovarian cancer
CC antigens) genes, and AAE12623-AAE12629 represent the proteins they
CC encode. AAE12630-AAE12638 represent human ovarian cancer antigenic
CC fragments or variants. Ovarian cancer antigens and their corresponding
CC DNAs are used in the prevention, diagnosis and treatment of diseases
CC associated with their inappropriate expression. These disorders include
CC proliferative disorders, cancer, tumours, respiratory system disorders,
CC asthma, haematopoietic disorders, diseases of the immune system, AIDS,
CC skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
CC inflammation, allergies, neurological disorders (e.g., Alzheimer's
CC disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, endocrine disorders and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, to identify their cognate ligands or binding
CC partners, in chemotaxis and can be used as a food additive. Antibodies
CC specific for a protein of the invention can be used in alleviating
CC symptoms associated with the disorders mentioned above and in diagnostic
CC immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
CC the invention is used in gene therapy. The present sequence represents a
CC human ovarian cancer antigenic fragment of the invention

XX Sequence 526 AA;

Query Match 100.0%; Score 104; DB 4; Length 526;
Best Local Similarity 100.0%; Pred. No. 1, 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNRPITGNS 20
DB 448 SSVLDGYSPPNRPITGNS 467

RESULT 14

ID ADS94302 standard; protein; 748 AA.

AC ADS94302;

DT 02-DEC-2004 (first entry)

XX CA 125/0772P 3-repeat amino acid sequence SEQ ID NO:1.

XX antibody; antigen-binding antibody fragment;

XX cell-associated CA 125/0772P; monoclonal antibody; cytostatic;

XX immunostimulant; mediator of lysis; tumour; cell proliferative disorder;

XX cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;

XX ovarian cancer.

XX Synthetic.

XX WO2004035537-A2.
 XX
 XX 29-APR-2004.
 PD
 XX
 XX 15-OCT-2003; 2003WO-US032945.
 PF
 XX
 XX 16-OCT-2002; 2002US-0418828P.
 PR
 XX 10-JUL-2003; 2003US-0485986P.
 XX
 XX (EURO-) EUROCELTIQUE SA.
 PA
 XX Albione EF, Solitis DA;
 PI
 XX WPI; 2004-357171/33.
 DR
 XX
 XX Novel isolated antibody, or antigen-binding antibody fragment binding
 PT with cell-associated CA 125/O772P polypeptide relative to shed CA
 PT 125/O772 polypeptide, useful for ameliorating cervical or ovarian cancer.
 XX
 XX Example; SEQ ID NO 1, 153bp; English.

The present invention describes an isolated antibody, or an antigen-binding antibody fragment (I), that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772 polypeptide. Also described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2) a monoclonal antibody that competes with binding of (II); (3) a hybridoma as deposited in (II); (4) an isolated nucleic acid molecule (III) comprising a nucleotide sequence that encodes a variable chain region of (II); (5) a pharmaceutical composition comprising an antibody or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (6) a pharmaceutical composition comprising a monoclonal antibody or an antigen-binding monoclonal antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (7) an article of manufacture (IV) comprising packaging material and a composition comprising an antibody, or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (8) a fusion polypeptide (V) comprising an antibody, or an antigen-binding antibody fragment, which preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P operably linked to a heterologous agent; (9) ameliorating (M1) a symptom of a CA 125/O772P-related disorder; (10) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1, 725.1, 8G9, 7F10, 8A1, 8AC3, 15C9, 8B3, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9, 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding antibody fragment; (11) an antibody or antigen binding antibody fragment that competes with (VI); and (12) a pharmaceutical composition comprising (VI) and a carrier. (I) has cytostatic activity, and can be used as an immunostimulant and a mediator of lysis of positive tumour cell. (I) is useful for ameliorating a symptom of a CA 125/O77P-related disorder which is a cell proliferative disorder such as cancer, cervical or uterine cancer, breast or lung cancer or ovarian cancer. (V) is useful diagnostically for monitoring the development or progression of cancer or tumour as part of clinical testing procedure. The present sequence represents the CA 125/O772P 3-repeat amino acid sequence, which is used in the exemplification of the present invention.

XX Sequence 748 AA;
 SQ
 Query Match 100.0%; Score 104; DB 8; Length 748;
 Beet Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVLVDSPPNNEPLTGS 20
 |||||
 DB 691 SSVLVDSPPNNEPLTGS 710

RESULT 15
 ADS94303

ID ADS94303 standard; protein; 809 AA.
 XX
 XX ADS94303;
 AC
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX
 XX CA 125/O772P 3-repeat TM amino acid sequence SEQ ID NO:2.
 DE
 XX
 XX antibody; antigen-binding antibody fragment;
 KW cell-associated CA 125/O772P; monoclonal antibody; cytostatic;
 KW immunostimulant; mediator of lysis; tumour; cell proliferative disorder;
 KW cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;
 KW ovarian cancer.
 KM
 XX Synthetic.
 OS
 XX
 XX WO2004035537-A2.
 PN
 XX
 XX 29-APR-2004.
 PD
 XX
 XX 15-OCT-2003; 2003WO-US032945.
 PF
 XX
 XX 16-OCT-2002; 2002US-0418828P.
 PR
 XX 10-JUL-2003; 2003US-0485986P.
 XX
 XX (EURO-) EUROCELTIQUE SA.
 PA
 XX Albione EF, Solitis DA;
 PI
 XX WPI; 2004-357171/33.
 DR
 XX
 XX Novel isolated antibody, or antigen-binding antibody fragment binding
 PT with cell-associated CA 125/O772P polypeptide relative to shed CA
 PT 125/O772 polypeptide, useful for ameliorating cervical or ovarian cancer.
 XX
 XX Example; SEQ ID NO 2; 153bp; English.

The present invention describes an isolated antibody, or an antigen-binding antibody fragment (I), that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772 polypeptide. Also described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2) a monoclonal antibody that competes with binding of (II); (3) a hybridoma as deposited in (II); (4) an isolated nucleic acid molecule (III) comprising a nucleotide sequence that encodes a variable chain region of (II); (5) a pharmaceutical composition comprising an antibody or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (6) a pharmaceutical composition comprising a monoclonal antibody or an antigen-binding monoclonal antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P operably linked to a heterologous agent; (9) ameliorating (M1) a symptom of a CA 125/O772P-related disorder; (10) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1, 725.1, 8G9, 7F10, 8A1, 8AC3, 15C9, 8B3, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9, 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding antibody fragment; (11) an antibody or antigen binding antibody fragment that competes with (VI); and (12) a pharmaceutical composition comprising (VI) and a carrier. (I) has cytostatic activity, and can be used as an immunostimulant and a mediator of lysis of positive tumour cell. (I) is useful for ameliorating a symptom of a CA 125/O77P-related disorder which is a cell proliferative disorder such as cancer, cervical or uterine cancer, breast or lung cancer or ovarian cancer. (V) is useful diagnostically for monitoring the development or progression of cancer or tumour as part of clinical testing procedure. The present sequence represents the CA 125/O772P 3-repeat TM amino acid sequence, which is

CC used in the exemplification of the present invention.

XX Sequence 809 AA;

Query Match 100.0%; Score 104; DB 8; Length 809;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLYDGYSPNREPLTGNS 20
 DB 691 SSVLYDGYSPNREPLTGNS 710

Search completed: October 27, 2005, 18:44:23
 Job time : 73.2093 secs

This page blank (us9:9)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using ew model

Run on: October 27, 2005, 17:03:42 ; Search time 17.3953 Seconds
(without alignments)
85.826 Million cell updates/sec

Title: US-10-612-090-14
Perfect score: 104
Sequence: 1 SSVLDVGYSPNNEPLTGS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/SB_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	94.2	438	4	US-09-404-879A-390
2	98	94.2	438	4	US-09-667-857-390
3	98	94.2	833	4	US-09-404-879A-389
4	98	94.2	833	4	US-09-667-857-389
5	98	94.2	914	4	US-09-404-879A-312
6	98	94.2	914	4	US-09-338-933-312
7	98	94.2	914	4	US-09-667-857-312
8	50	48.1	3519	3	US-09-428-517-4
9	47	45.2	2142	4	US-09-540-236-3459
10	45	43.3	513	1	US-08-464-340A-2
11	45	43.3	513	4	US-09-949-016-10288
12	45	43.3	513	5	PCT-US94-0849A-2
13	45	43.3	811	4	US-09-489-039A-13839
14	45	43.3	1048	4	US-09-171-699-10
15	44	42.3	1715	4	US-09-696-115B-17
16	43	41.8	299	4	US-09-902-540-13589
17	43	41.3	391	2	US-08-759-581B-13
18	43	41.3	391	3	US-09-304-711-13
19	43	41.3	391	4	US-09-173-281-13
20	43	41.3	504	4	US-09-489-039A-11966
21	43	41.3	575	4	US-09-489-039A-8630
22	43	41.3	994	4	US-09-521-511C-13
23	43	41.3	1346	3	US-09-320-878-4
24	43	41.3	1346	3	US-09-105-537-37
25	43	41.3	1346	4	US-09-141-908-5
26	43	41.3	1346	4	US-09-657-440-4
27	43	41.3	11877	3	US-09-105-537-6

28	42.5	40.9	428	4	US-09-538-092-531	Sequence 531, App
29	42	40.4	328	4	US-09-902-540-10796	Sequence 10796, A
30	42	40.4	422	4	US-09-831-630-10	Sequence 10, Appl
31	42	40.4	721	4	US-09-328-352-7781	Sequence 7781, Ap
32	41.5	39.9	318	4	US-09-328-352-6869	Sequence 6869, Ap
33	41	39.4	201	3	US-09-015-734-12	Sequence 12, Appl
34	41	39.4	201	4	US-09-515-311-12	Sequence 14310, A
35	41	39.4	213	4	US-09-489-039A-14310	Sequence 7, Appl1
36	41	39.4	236	3	US-09-015-734-7	Sequence 2, Appl1
37	41	39.4	236	4	US-09-515-311-7	Sequence 2, Appl1
38	41	39.4	255	3	US-09-015-734-2	Sequence 27314, A
39	41	39.4	255	4	US-09-515-311-2	Sequence 26329, A
40	41	39.4	468	4	US-09-248-796A-27314	Sequence 14460, A
41	41	39.4	641	4	US-09-252-991A-26329	Sequence 19979, A
42	41	39.4	655	4	US-09-248-796A-14460	Sequence 8938, Ap
43	41	39.4	746	4	US-09-248-796A-19979	Sequence 4591, Ap
44	41	39.4	922	4	US-09-489-039A-8938	
45	41	39.4	929	4	US-09-328-352-4591	

ALIGNMENTS

RESULT 1
US-09-404-879A-390
; Sequence 390, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 390
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-390

Query Match 94.2%; Score 98; DB 4; Length 438;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVLDVGYSPNNEPLTGS 20
DB 360 SSVLDVGYSPNNEPLTGS 379

RESULT 2
US-09-667-857-390
; Sequence 390, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 390
LENGTH: 438
TYPE: PRT
ORGANISM: Homo sapiens
US-09-667-857-390

Query Match 94.2%; Score 98; DB 4; Length 438;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNNEPLTGS 20
DB 360 SSVLDGTFPNNNEPLTGS 379

RESULT 3
US-09-404-879A-389
Sequence 389, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-09-404-879A-389

Query Match 94.2%; Score 98; DB 4; Length 833;
Best Local Similarity 95.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNNEPLTGS 20
DB 755 SSVLDGTFPNNNEPLTGS 774

RESULT 4
US-09-667-857-389
Sequence 389, Application US/09667857
Patent No. 6699664
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Steven W.
APPLICANT: Fanger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C5
CURRENT APPLICATION NUMBER: US/09/667,857
CURRENT FILING DATE: 2000-09-20
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-09-667-857-389

Query Match 94.2%; Score 98; DB 4; Length 833;
Best Local Similarity 95.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNNEPLTGS 20
DB 755 SSVLDGTFPNNNEPLTGS 774

RESULT 5
US-09-404-879A-312
Sequence 312, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 312
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapien
US-09-404-879A-312

Query Match 94.2%; Score 98; DB 4; Length 914;
Best Local Similarity 95.0%; Pred. No. 4.2e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNNEPLTGS 20
DB 836 SSVLDGTFPNNNEPLTGS 855

RESULT 6
US-09-338-933-312
Sequence 312, Application US/09338933
Patent No. 6468931
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer Lynn
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
FILE REFERENCE: 210121.462C1
CURRENT APPLICATION NUMBER: US/09/338,933
CURRENT FILING DATE: 1999-06-23
NUMBER OF SEQ ID NOS: 312
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 312
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapien
US-09-338-933-312

Query Match 94.2%; Score 98; DB 4; Length 914;
Best Local Similarity 95.0%; Pred. No. 4.2e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVLDGYSPPNNEPLTGS 20
DB 836 SSVLDGTFPNNNEPLTGS 855

RESULT 7
US-09-667-857-312
Sequence 312, Application US/09667857
Patent No. 6699664
GENERAL INFORMATION:


```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Rling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.46205
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-667-857-312

```

```

Query Match          94.2%; Score 98; DB 4; Length 914;
Best Local Similarity 95.0%; Pred. No. 4.2e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 SSVLDGSPNRPNEPLTGN 20
DB      836 SSVLDGYFPNRPNEPLTGN 855

```

```

RESULT 8
US-09-428-517-4
; Sequence 4, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Beljach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-4

```

```

Query Match          48.1%; Score 50; DB 3; Length 3519;
Best Local Similarity 62.5%; Pred. No. 87;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      3 VLVYDGYSPNRPNEPLTG 18
DB      3401 VLVYDVPYPGROPEVFG 3416

```

```

RESULT 9
US-09-540-236-3459
; Sequence 3459, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:

```

```

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3459
; LENGTH: 2142
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-3459

```

```

Query Match          45.2%; Score 47; DB 4; Length 2142;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      2 SSVLDGSPNRPNEPLTGN 19
DB      105 STAIGGKFPNRPQDDNN 122

```

```

RESULT 10
US-08-464-340A-2
; Sequence 2, Application US/08464340A
; Patent No. 5710019
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Potassium Channel 1 and 2 Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,340A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/08449
; FILING DATE: 28 JUL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-415
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-464-340A-2

```

```

Query Match          43.3%; Score 45; DB 1; Length 513;
Best Local Similarity 47.4%; Pred. No. 57;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      2 SSVLDGSPNRPNEPLTGN 20
DB      315 TLVVDGAAAGRRKPGAGNS 333

```

```

RESULT 11
US-09-949-016-10288
; Sequence 10288, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10288
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10288

Query Match          43.3% Score 45; DB 4; Length 513;
Best Local Similarity 47.4% Pred. No. 57;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      2 SVLVGYSFNRNEPLTGN 20
DB      315 TLVVDGAAAGRRKPGAGNS 333

RESULT 12
PCT-US94-08449A-2
; Sequence 2, Application PC/TUS9408449A
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Potassium Channel Protein 1 and 2
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/08449A
; FILING DATE: SUBMITTED HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 AMINO ACIDS

```

```

; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
PCT-US94-08449A-2

Query Match          43.3% Score 45; DB 5; Length 513;
Best Local Similarity 47.4% Pred. No. 57;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      2 SVLVGYSFNRNEPLTGN 20
DB      315 TLVVDGAAAGRRKPGAGNS 333

RESULT 13
US-09-489-039A-13839
; Sequence 13839, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13839
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13839

Query Match          43.3% Score 45; DB 4; Length 811;
Best Local Similarity 66.7% Pred. No. 97;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      8 YSPNRNEPLTGN 19
DB      750 YSPNINIPINGN 761

RESULT 14
US-09-171-699-10
; Sequence 10, Application US/09171699
; Patent No. 6448389
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; ADDRESSEE: Gonczol, Eva
; ADDRESSEE: Berencsi, Klara
; ADDRESSEE: Kari, Csaba
; TITLE OF INVENTION: NO. 6448389el Cytomegalovirus DNA Constructs and
; Uses therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 60/015,717
FILING DATE: 23-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kodroff, Cathy A.
REGISTRATION NUMBER: 33,980
REFERENCE/DOCKET NUMBER: WST66APCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1048 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-171-699-10

Query Match 43.3%; Score 45; DB 4; Length 1048;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DGSPNNRNP 15
DB 537 DGYPNRDP 546

RESULT 15
US-09-696-115B-17
Sequence 17, Application US/09696115B
Patent No. 6660497
GENERAL INFORMATION:
APPLICANT: Bulla, Lee
APPLICANT: Candae, Mehmet
TITLE OF INVENTION: Pectinophora gossypiella (Pink Bollworm) Bacillus thuringiensis
TITLE OF INVENTION: toxin receptor BT-R2
FILE REFERENCE: 52418-2003.00
CURRENT APPLICATION NUMBER: US/09/696,115B
CURRENT FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/161,564
PRIOR FILING DATE: 1999-10-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 1715
TYPE: PRT
ORGANISM: B. mori
US-09-696-115B-17

Query Match 42.3%; Score 44; DB 4; Length 1715;
Best Local Similarity 61.5%; Pred. No. 3.4e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 YSPNNEPLTGN 20
DB 1028 YAPDRDEPDND 1040

Search completed: October 27, 2005, 17:34:31
Job time: 18.3953 secs

This Page Blank (1894)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:30:08 ; Search time 64 Seconds
(without alignments)
130.638 Million cell updates/sec

Title: US-10-612-090-14
Perfect score: 104
Sequence: 1 SSVLVGYSFNRNRLTGNS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues
Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US10E_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubppa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	20	15	US-10-612-090-14
2	104	100.0	108	15	US-10-612-090-2
3	104	100.0	108	15	US-10-612-090-13
4	104	100.0	174	15	US-10-383-368-4
5	104	100.0	284	10	US-09-965-738-300
6	104	100.0	438	9	US-09-884-441-483
7	104	100.0	438	10	US-09-907-969-483
8	104	100.0	438	14	US-10-198-053-483
9	104	100.0	438	15	US-10-333-900-19
10	104	100.0	438	17	US-10-860-790-483
11	104	100.0	439	10	US-09-965-738-148

12	104	100.0	456	18	US-10-858-412-226	Sequence 226, App
13	104	100.0	526	15	US-10-333-900-30	Sequence 30, Appl
14	104	100.0	545	15	US-10-243-243A-4	Sequence 4, Appl
15	104	100.0	583	14	US-10-142-515-4	Sequence 4, Appl
16	104	100.0	748	17	US-10-687-035-1	Sequence 1, Appl
17	104	100.0	809	17	US-10-687-035-2	Sequence 2, Appl
18	104	100.0	1148	9	US-09-884-441-458	Sequence 458, App
19	104	100.0	1148	9	US-09-884-441-479	Sequence 479, App
20	104	100.0	1148	10	US-09-907-969-458	Sequence 458, App
21	104	100.0	1148	10	US-09-907-969-479	Sequence 479, App
22	104	100.0	1148	10	US-09-932-419-2	Sequence 2, Appl
23	104	100.0	1148	10	US-09-965-738-48	Sequence 48, Appl
24	104	100.0	1148	10	US-09-827-271-458	Sequence 458, App
25	104	100.0	1148	14	US-10-097-340-216	Sequence 216, App
26	104	100.0	1148	14	US-10-198-053-458	Sequence 458, App
27	104	100.0	1148	14	US-10-198-053-479	Sequence 479, App
28	104	100.0	1148	15	US-10-257-021-56	Sequence 56, Appl
29	104	100.0	1148	15	US-10-383-368-2	Sequence 2, Appl
30	104	100.0	1148	16	US-10-734-564-131	Sequence 131, App
31	104	100.0	1148	17	US-10-860-790-458	Sequence 458, App
32	104	100.0	1148	17	US-10-860-790-479	Sequence 479, App
33	104	100.0	1148	17	US-10-936-626-129	Sequence 129, App
34	104	100.0	1148	17	US-10-938-061-129	Sequence 129, App
35	104	100.0	1148	20	US-11-050-926-216	Sequence 216, App
36	104	100.0	1156	9	US-09-884-441-459	Sequence 459, App
37	104	100.0	1156	10	US-09-907-969-459	Sequence 459, App
38	104	100.0	1156	10	US-09-827-271-459	Sequence 459, App
39	104	100.0	1156	14	US-10-198-053-459	Sequence 459, App
40	104	100.0	1156	17	US-10-860-790-459	Sequence 459, App
41	104	100.0	1889	14	US-10-142-515-5	Sequence 5, Appl
42	104	100.0	1890	14	US-10-097-340-217	Sequence 217, App
43	104	100.0	1890	15	US-10-245-871-314	Sequence 314, App
44	104	100.0	1890	15	US-10-253-266-314	Sequence 314, App
45	104	100.0	1890	20	US-11-050-926-217	Sequence 217, App

ALIGNMENTS

RESULT 1
US-10-612-090-14
; Sequence 14, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-14

Query Match 100.0%; Score 104; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLVGYSFNRNRLTGNS 20
DB 1 SSVLVGYSFNRNRLTGNS 20

RESULT 2
US-10-612-090-2
; Sequence 2, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:

```

; APPLICANT: ImmunoGen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-2

Query Match          100.0%; Score 104; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 6,6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVLDGYSPPNNEPLTGNS 20
DB      86 SSVLDGYSPPNNEPLTGNS 105

RESULT 3
US-10-612-090-13
; Sequence 13, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: ImmunoGen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(1)
; OTHER INFORMATION: Glutathione S-transferase fusion site
US-10-612-090-13

Query Match          100.0%; Score 104; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 6,6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVLDGYSPPNNEPLTGNS 20
DB      86 SSVLDGYSPPNNEPLTGNS 105

RESULT 4
US-10-383-368-4
; Sequence 4, Application US/10383368
; Publication No. US20040002091A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul B.
; APPLICANT: Ebner, Reinhard
; APPLICANT: Scrovel, Jeffrey W.
; TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
; FILE REFERENCE: 689290-123
; CURRENT APPLICATION NUMBER: US/10/383,368
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US 60/362,527
; PRIOR FILING DATE: 2002-03-07
```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-383-368-4

Query Match          100.0%; Score 104; DB 15; Length 174;
Best Local Similarity 100.0%; Pred. No. 1,1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVLDGYSPPNNEPLTGNS 20
DB      96 SSVLDGYSPPNNEPLTGNS 115

RESULT 5
US-09-965-738-300
; Sequence 300, Application US/09965738
; Publication No. US2003014367A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy
; TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic ar
; FILE REFERENCE: 40715-258841
; CURRENT APPLICATION NUMBER: US/09/965,738
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/284,175
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 300
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-738-300

Query Match          100.0%; Score 104; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVLDGYSPPNNEPLTGNS 20
DB      206 SSVLDGYSPPNNEPLTGNS 225

RESULT 6
US-09-884-441-483
; Sequence 483, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121,462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 483
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-483

Query Match          100.0%; Score 104; DB 9; Length 438;
Best Local Similarity 100.0%; Pred. No. 3,3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVLDGYSPPNNEPLTGNS 20
```

Db 360 SSVLDVGYSPNNEPITGNS 379

RESULT 7

US-09-907-969-483
; Sequence 483, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 438
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-907-969-483

Query Match 100.0%; Score 104; DB 10; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVLDVGYSPNNEPITGNS 20
|||||
Db 360 SSVLDVGYSPNNEPITGNS 379

RESULT 8
US-10-198-053-483
; Sequence 483, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 438
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-198-053-483

Query Match 100.0%; Score 104; DB 14; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVLDVGYSPNNEPITGNS 20
|||||
Db 360 SSVLDVGYSPNNEPITGNS 379

RESULT 9

US-10-333-900-19
; Sequence 19, Application US/10333900
; Publication No. US20040010121A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 7 Human Ovarian and Ovarian Cancer Associated Proteins
; FILE REFERENCE: PA006PCT
; CURRENT APPLICATION NUMBER: US/10/333,900
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/227,009
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/190,076
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 438
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-333-900-19

Query Match 100.0%; Score 104; DB 15; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVLDVGYSPNNEPITGNS 20
|||||
Db 360 SSVLDVGYSPNNEPITGNS 379

RESULT 10
US-10-860-790-483
; Sequence 483, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 438
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-860-790-483

Query Match 100.0%; Score 104; DB 17; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVLDVGYSPNNEPITGNS 20
|||||
Db 360 SSVLDVGYSPNNEPITGNS 379

RESULT 11
US-09-965-738-148
; Sequence 148, Application US/09965738
; Publication No. US2003014367A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy
; TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic ar
; FILE REFERENCE: 40715-258841
; CURRENT APPLICATION NUMBER: US/09/965,738
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/284,175

;; PRIOR FILING DATE: 2001-04-17
;; NUMBER OF SEQ ID NOS: 306
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 148
;; LENGTH: 439
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-738-148

Query Match 100.0%; Score 104; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDVGYSPNREPLTGS 20
|||
Db 361 SSVLDVGYSPNREPLTGS 380

RESULT 12
US-10-858-412-226
; Sequence 226, Application US/10858412
; Publication No. US20050214781A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Turner, Leah
; APPLICANT: Sun, Yongming
; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Ovarian Spe
; FILE REFERENCE: DEX-0488
; CURRENT APPLICATION NUMBER: US/10/858,412
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/474,854
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 226
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-858-412-226

Query Match 100.0%; Score 104; DB 18; Length 456;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDVGYSPNREPLTGS 20
|||
Db 378 SSVLDVGYSPNREPLTGS 397

RESULT 13
US-10-333-900-30
; Sequence 30, Application US/10333900
; Publication No. US20040010121A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 7 Human Ovarian and Ovarian Cancer Associated Proteins
; FILE REFERENCE: PA006PCT
; CURRENT APPLICATION NUMBER: US/10/333,900
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/227,009
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/190,076
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (116)

;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-333-900-30

Query Match 100.0%; Score 104; DB 15; Length 526;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDVGYSPNREPLTGS 20
|||
Db 448 SSVLDVGYSPNREPLTGS 467

RESULT 14
US-10-243-243A-4
; Sequence 4, Application US/10243243A
; Publication No. US20030104442A1
; GENERAL INFORMATION:
; APPLICANT: Lloyd, Kenneth O.
; APPLICANT: Yin, Beatrice W.T.
; TITLE OF INVENTION: Nucleic Acid Sequence Encoding Ovarian Antigen, CA125, and Uses Th
; FILE REFERENCE: 649-B
; CURRENT APPLICATION NUMBER: US/10/243,243A
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/142,515
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: PCT/US02/14768
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US 60/290,480
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(545)
; OTHER INFORMATION: Deduced amino acid sequence of CA125/MUC16A (B4), Part Three
US-10-243-243A-4

Query Match 100.0%; Score 104; DB 14; Length 545;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLDVGYSPNREPLTGS 20
|||
Db 467 SSVLDVGYSPNREPLTGS 486

RESULT 15
US-10-142-515-4
; Sequence 4, Application US/10142515
; Publication No. US20030078399A1
; GENERAL INFORMATION:
; APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
; APPLICANT: Lloyd, Kenneth O.
; APPLICANT: Yin, Beatrice W.T.
; TITLE OF INVENTION: Nucleic Acid Sequence Encoding Ovarian Antigen, CA125, and Uses Th
; FILE REFERENCE: 649-A-US
; CURRENT APPLICATION NUMBER: US/10/142,515
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/290,480
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Human Being
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(583)

; OTHER INFORMATION: Deduced amino acid sequence of CA125/MUC16 (B4), Section three
US-10-142-515-4

Query Match 100.0%; Score 104; DB 14; Length 583;
Best Local Similarity 100.0%; Pred. No. 4.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVLVGDGSPNRPPLTGNS 20
|||
Db 505 SSVLVGDGSPNRPPLTGNS 524

Search completed: October 27, 2005, 18:31:45
Job time : 65 secs

SECRET

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 18:13:49 ; Search time 12.3721 Seconds
(without alignments)
155.538 Million cell updates/sec

Title: US-10-612-090-14

Perfect score: 104

Sequence: 1 SSVLVNDGSPNRPNEPLTGN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

PIR 79: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	48.1	3519	2	S43048	polyketide synthas
2	49	47.1	1666	1	C3GP	complement C3 prec
3	47	45.2	407	2	S49890	regulatory protein
4	46	44.2	1018	2	JC5799	alpha-D-mannosidas
5	46	44.2	1076	2	T26044	hypothetical prote
6	45	43.3	107	2	T46746	ubiquinol-cytochro
7	45	43.3	513	2	JC5920	potassium channel
8	45	43.3	715	2	AE1021	formate dehydrogen
9	45	43.3	1048	1	XPBEA9	large structural p
10	45	43.3	1309	2	T39379	sexual differential
11	45	42.3	2214	2	T16305	hypothetical prote
12	44	42.3	186	2	C81903	hypothetical prote
13	44	42.3	253	2	D81118	conserved hypotet
14	44	42.3	445	2	AD2023	carboxyl-terminal
15	44	42.3	748	2	T35818	probable integral
16	44	42.3	1015	2	T32186	hypothetical prote
17	44	42.3	1355	2	T28715	hypothetical prote
18	44	42.3	1715	2	JEO128	Bombyx mori recept
19	44	41.8	1006	2	AD2195	hypothetical prote
20	43	41.3	367	2	T23983	hypothetical prote
21	43	41.3	426	1	DEBCFS	conserved hypotet
22	43	41.3	715	1	DEBCFS	formate dehydrogen
23	43	41.3	715	2	E91261	formate dehydrogen
24	43	41.3	715	2	A86102	formate dehydrogen
25	43	41.3	1346	2	T17412	polyketide synthas
26	42.5	40.9	428	2	S61623	hypothetical prote
27	42.5	40.9	1089	2	T21582	hypothetical prote
28	42	40.4	62	2	JQ0208	hypothetical 6.6k
29	42	40.4	62	2	S07166	hypothetical prote

30	42	40.4	62	2	S58538	hypothetical prote
31	42	40.4	70	2	T04410	hypothetical prote
32	42	40.4	78	2	C95976	hypothetical prote
33	42	40.4	295	2	S58850	homeotic protein a
34	42	40.4	320	2	T36057	hypothetical prote
35	42	40.4	474	2	C87556	hypothetical prote
36	42	40.4	507	1	A30828	steroid 17alpha-mo
37	42	40.4	582	2	S40176	Exor protein - Rhl
38	42	40.4	613	2	F81334	hypothetical prote
39	42	40.4	648	2	B40727	S-M checkpoint con
40	42	40.4	792	2	F83304	probable restricti
41	42	40.4	1663	1	C3RT	complement C3 prec
42	42	40.4	1778	2	T50074	probable nucleopor
43	41.5	39.9	225	2	A86903	hypothetical prote
44	41.5	39.9	735	2	G71203	probable translati
45	41	39.4	144	2	T29958	hypothetical prote

ALIGNMENTS

RESULT 1

S43048 polyketide synthase type I - Streptomyces antibioticus
N:Contains: acyl carrier protein; acyltransferase; ketosacyl synthase; ketoreductase; thioe
C:Species: Streptomyces antibioticus

C>Date: 13-Jan-1995 #sequence_revision 06-Dec-1996 #text_change 16-Aug-2004

C/Accession: S43048; S41729

R:Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.

submitted to the EMBL Data Library, February 1993

A/Reference number: S43048

A/Accession: S43048

A/Molecule type: DNA

A/Residues: 1-3519 <SWA>

A/Cross-references: UNIPROT:Q07017; EMBL:L09654; NID:G153407; PID:AAA19695.1; PID:G15340

R:Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.

Mol. Gen. Genet. 242, 358-362, 1994

A/Title: Characterisation of a Streptomyces antibioticus gene encoding a type I polyketic

A/Reference number: S41729; NID:94150470; PMID:8107683

A/Accession: S41729

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1683-3238, 3273-3303, 'K', 3305-3407, 'T', 3409-3462, 'Y', 3464-3516, 'E', 3518-3519

A/Cross-references: EMBL:L09654

C/Genes:

C/Start codon: GTG

C/Keywords: antibiotic biosynthesis; carrier protein

F:59-462/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F:570-851/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F:1202-1181/Domain: short-chain alcohol dehydrogenase homology <SAD1>

F:1489-1560/Domain: acyl carrier protein homology <ACPI>

F:1708-2111/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F:2221-2502/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>

F:2858-3037/Domain: short-chain alcohol dehydrogenase homology <SAD2>

F:3143-3214/Domain: acyl carrier protein homology <ACP2>

F:3305-3500/Domain: oleoyl-[acyl-carrier-protein] hydrolyase homology <ACPH>

Query Match 48.1%; Score 50; DB 2; Length 3519;

Best Local Similarity 62.5%; Pred. No. 40;

Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 VLVNDGSPNRPNEPLNG 18

DB 3401 VLVNDVPPGRQEPVFG 3416

RESULT 2

C3GP Complement C3 precursor - guinea pig

N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit; (

C/Species: Cavia porcellus (guinea pig)

C>Date: 07-Feb-1992 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004

C/Accession: A37156; S03375; A20342; D20342; C20342; A31222

R:Auerbach, H.S.; Burger, R.; Dodds, A.; Colten, H.R.
 J. Clin. Invest. 86, 96-106, 1990
 A>Title: Molecular basis of complement C3 deficiency in guinea pigs.
 A:Reference number: A37156; MUID:90307998; PMID:1973176
 A:Accession: A37156
 A:Molecule type: mRNA
 A:Residues: 1-166 <AUB>
 A:Cross-references: UNIPROT:P12387; GB:M34054; NID:9191262; PIDN:AAA7038.1; PID:9305335
 R:Gerard, N.P.; Lively, M.O.; Gerard, C.
 Protein Seq. Data Anal. 1, 473-478, 1988
 A>Title: Amino acid sequence of guinea pig C3a anaphylatoxin.
 A:Reference number: S03375; MUID:89113342; PMID:3064079
 A:Accession: S03375
 A:Molecule type: Protein
 A:Residues: 676-730, 'N', '732-752' <GER>
 A:Experimental source: complement-activated guinea pig serum
 A>Note: form isolated is inactive C3a anaphylatoxin and is missing the carboxyl-terminal
 R:Thomae, M.L.; Tack, B.F.
 Biochemistry 22, 942-947, 1983
 A>Title: Identification and alignment of a thiol ester site in the third component of gu
 A:Reference number: A50479; MUID:83178889; PMID:6838833
 A:Accession: A20342
 A:Molecule type: protein
 A:Residues: 676-687 <TH1>
 A:Accession: D20342
 A:Molecule type: protein
 A:Residues: 993-1012, 1014-1017, 'E', '1019-1030', 'Y' <TH2>
 R:Goldberger, G.; Thomas, M.L.; Tack, B.F.; Williams, J.; Colten, H.R.; Abraham, G.N.
 J. Biol. Chem. 256, 12617-12619, 1981
 A>Title: NH2-terminal structure and cleavage of guinea pig pro-C3, the precursor of the
 A:Reference number: A20342; MUID:82075767; PMID:6458605
 A:Accession: C20342
 A:Molecule type: protein
 A:Residues: 23-38 <GOB>
 C:Comment: Complement C3 contains two chains, formed by removal of four residues and 11f
 alternative-complement pathway; releases the C3a anaphylatoxin from the amino end of t
 inactive-complement-pathway C3/C5 convertase.
 C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
 C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
 classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
 C:Comment: The major site of synthesis of this plasma protein is the liver.
 C:Superfamily: alpha-2-macroglobulin
 C:Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
 F:1-22/Domain: signal sequence #status predicted <SID>
 F:23-671/Product: complement C3 and C3b beta chain #status predicted <C3BB>
 F:23-671, 676-1666/Product: complement C3 #status predicted <CC3>
 F:23-671, 754-1666/Product: complement C3b #status predicted <C3B>
 F:676-1666/Product: complement C3 alpha chain #status predicted <C3A>
 F:676-753/Product: C3a anaphylatoxin #status predicted <C3a>
 F:754-1666/Product: complement C3b alpha' chain #status predicted <C3BA>
 F:951-1308/Product: C3d fragment #status predicted <CD3>
 F:1007-1308/Product: C3d fragment #status predicted <CD3>
 F:1429-1461/Region: properdin binding
 F:557-821, 630-666, 698-725, 699-732, 712-733, 878-1517, 1106-1163, 1363-1493, 1394-1462, 1510-15
 F:753-754/Cleavage site: Arg-Ser (C3 convertase) #status predicted
 F:944, 1620/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:1015-1018/Cross-link: thioester (Cys-Gln) #status experimental
 F:1308-1309/Cleavage site: Arg-Ser (complement factor I) #status predicted
 F:1325-1326/Cleavage site: Arg-Ser (complement factor I) #status predicted

Query Match 47.1%; Score 49; DB 1; Length 1666;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Db 305 VLDGVQSPRRPALVKS 322

RESULT 3
 S49890
 regulatory protein RPI1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein Y18277.10c; protein Y1L19c

C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
 A:Accession: S49890; A41218
 R:Hamlyn, N.; Churcher, C.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: S49881
 A:Accession: S49890
 A:Molecule type: DNA
 A:Residues: 1-407 <HAM>
 A:Cross-references: UNIPROT:P23250; GB:Z47047; EMBL:Z46833; NID:9603997; PID:9763227; MII
 R:Kim, J.H.; Powers, S.
 Mol. Cell. Biol. 11, 3894-3904, 1991
 A>Title: Overexpression of RPI1, a novel inhibitor of the yeast Ras-cyclic AMP pathway, c
 A:Reference number: A41218; MUID:91304380; PMID:1649384
 A:Accession: A41218
 A:Molecule type: DNA
 A:Residues: 1-16, 'K', '38-70', 'P', '74-90', 'V', '92-169', '175-239', 'SNS', '243-247', 'NNSNNN', '248-407' <
 A:Cross-references: GB:S42682; NID:91679961; PIDN:ABB39643.1; PID:9233044
 C:Genetics:
 A:Gene: SGD:RPI1
 A:Cross-references: SGD:S0001381; MIPS:Y1L19c
 A:Map position: 9L

Query Match 45.2%; Score 47; DB 2; Length 407;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 DGVSPNRNEPLTGN 19
 DB 26 YSPNINPTITSN 37

RESULT 4
 JCS799
 alpha-D-mannosidase (EC 3.2.1.-) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
 A:Accession: JCS799
 R:Hixmoco, S.; Tamba, M.; Kiuchi, S.; Jin, Y.Z.; Bannai, S.; Sugita, Y.; Dacheux, F.; De
 Blochem. Biophys. Res. Commun. 241, 439-445, 1997
 A>Title: Stage-specific expression of a mouse homologue of the porcine J35KDa alpha-D-mar
 A:Reference number: JCS799; MUID:98086569; PMID:9425289
 A:Accession: JCS799
 A:Molecule type: mRNA
 A:Residues: 1-1018 <HNR>
 A:Cross-references: UNIPROT:O54782; DBJ:AB006458; NID:92766160; PIDN:BA24266.1; PID:dl1
 A:Experimental source: testis
 C:Comment: This protein serves as a marker for the late stages of type A sperma togonia t
 C:Keywords: glycoprotein; glycosidase; hydrolase
 F:226,249,294,336,511,528,620,652,818/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 44.2%; Score 46; DB 2; Length 1018;
 Best Local Similarity 61.5%; Pred. No. 42;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DGVSPNRNEPLTG 18
 DB 244 DGVYPMSEPTVG 256

RESULT 5
 T26044
 hypothetical protein W01C8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 A:Accession: T26044
 R:Rhan, M.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans coemid W01C8.
 A:Reference number: Z20142
 A:Accession: T26044
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-1076 <NHA>
 A:Cross-references: EMBL:U41508; PIDN:AAA82623.1; CESP:W01C8.3
 C:Genetics:
 A:Gene: CESP:W01C8.3
 A:Introns: 59/3; 92/2; 157/3; 189/3; 220/2; 251/3; 275/2; 319/1; 374/3; 407/2

Query Match 44.2%; Score 46; DB 2; Length 1076;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LVVDGSPNRNEPLTNGS 14
 DB 837 IIDGYHPERNE 847

RESULT 6

ubiquitinol-cytochrome-c reductase (EC 1.10.2.2) chain VIII [Imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004

R:Labo:Hadju G.; Braun, H.P.; Romp, N.; Grivell, L.A.; Berden, J.A.; Schmitz, U.K.
 Biochem. J. 320, 769-775, 1996
 A:Title: Subunit VII of ubiquitinol-cytochrome-c oxidoreductase from Neurospora crassa is
 A:Reference number: 223142; MUID:97156980; PMID:9003361

A:Accession: T46746
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-107 <LMB>

A:Cross-references: UNIPROT:P48503; EMBL:U20790; NID:9687738; PIDN:AAC49654.1; PID:96877
 A:Experimental source: strain 74-OR23-1A; clone 5

C:Genetics:
 A:Gene: OCR8

C:Superfamily: Schistosacharomyces pombe ubiquitinol-cytochrome-c reductase chain VIII
 C:Keywords: electron transfer; membrane-associated complex; mitochondrion; oxidative pho

Query Match 43.3%; Score 45; DB 2; Length 107;
 Best Local Similarity 57.1%; Pred. No. 4.5;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 GYSNRPNEPLTNGS 20
 DB 41 GISPNRQNPICATGA 54

RESULT 7

JCS920
 potassium channel 2 - human

C:Species: Homo sapiens (man)
 C:Date: 28-Mar-1998 #sequence_revision 28-Mar-1998 #text_change 09-Jul-2004

C:Accession: JCS920
 R:Su, K.; Kyaw, H.; Fan, P.; Zeng, Z.; Shell, B.K.; Carter, K.C.; Li, Y.

Biochem. Biophys. Res. Commun. 242, 675-681, 1997
 A:Title: Isolation, characterization, and mapping of two human potassium channels.

A:Reference number: JCS919
 A:Accession: JCS920

A:Molecule type: mRNA
 A:Residues: 1-513 <SUA>

A:Cross-references: UNIPROT:Q9UIK4; GB:AF033383; NID:92739502; PIDN:AAC05635.1; PID:9273
 C:Comment: This plasma membrane protein has six transmembrane domains, and is involved in
 an external stimulus of neurotransmitters and neuropeptides.

C:Genetics:
 A:Gene: KH2

A:Map position: 20q13
 C:Superfamily: potassium channel protein drkl

Query Match 43.3%; Score 45; DB 2; Length 513;
 Best Local Similarity 47.4%; Pred. No. 28;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 SVLVDDGSPNRNEPLTNGS 20
 DB 315 TLVVDGAAAGRRKRGAGNS 333

RESULT 8

AE1021
 formate dehydrogenase (EC 1.2.1.2) - Salmonella enterica subsp. enterica serovar Typhi (

C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AE1021

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AE1021

A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-715 <PAR>
 A:Cross-references: UNIPROT:Q8Z1Q0; GB:AL513382; PIDN:CAD09270.1; PID:916505274; GSPDB:G

C:Genetics:
 A:Gene: STY4484

C:Superfamily: formate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 43.3%; Score 45; DB 2; Length 715;
 Best Local Similarity 47.4%; Pred. No. 41;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 SVLVDDGSPNRNEPLTNGS 20
 DB 256 SKVBEYTPESVSEIRGVGS 274

RESULT 9

XPBBA9
 large structural phosphoprotein p150 - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A29533; S09795
 R:Jahn, G.; Kouzarides, T.; Mach, M.; Scholl, B.C.; Plachter, B.; Traupe, B.; Preddie, E.

J. Virol. 61, 1358-1367, 1987
 A:Title: Map position and nucleotide sequence of the gene for the large structural phospho

A:Reference number: A29533; MUID:87198858; PMID:3033266
 A:Accession: A29533

A:Molecule type: DNA
 A:Accession: A29533

A:Residues: 1-1048 <JAH>
 A:Cross-references: UNIPROT:P08318; GB:M16022; NID:9330643; PIDN:AAA45992.1; PID:9330644

A:Gene: MS; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; i
 M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MUID:90269039; PMID:2161319
 A:Accession: S09795

A:Molecule type: DNA
 A:Residues: 1-1048 <CHB>

A:Cross-references: EMBL:X17403; NID:958591; PIDN:CAA35431.1; PID:959637
 C:Genetics:

A:Map position: 0.160-0.186
 C:Superfamily: human cytomegalovirus large structural phosphoprotein; large structural p

C:Keywords: phosphoprotein
 F;2-366/Domain: large structural phosphoprotein homology <CLS>

Query Match 43.3%; Score 45; DB 1; Length 1048;
 Best Local Similarity 70.0%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DGSPNRNEP 15
 DB 537 DGTPNRKDP 546

```

RESULT 10
T39379
sexual differentiation and meiosis protein ste20 - fission yeast (Schizosaccharomyces po
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39379; T46557
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, September 1995
A:Reference number: Z21849
A:Accession: T39379
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1309 <DEV>
A:Cross-References: UNIPROT:Q09743; EMBL:254140; PIDN:CAA90815.1; GSPDB:GN00066; SPDB:SE
A:Experimental source: strain 972h-, cosmid c12C2
R:Hiltl, N.; Baumann, D.; Edenharter, E.; Stalder, M.; Schweingruber, M.E.
submitted to the EMBL Data Library, February 1998
A:Description: A new gene, ste20, in fission yeast Schizosaccharomyces pombe - necessary
and pH regulation.
A:Reference number: Z23064
A:Accession: T46557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1309 <HIL>
A:Cross-References: EMBL:AJ23984; PIDN:CAA11758.1
C:Genetics:
A:Gene: ste16; SPBC12C2.02c; ste20
A:Map position: 1
A:Function:
A:Description: necessary for sexual development and also responsible for sensitivity, me

Query Match
Best Local Similarity 43.3%; Score 45; DB 2; Length 1309;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 DGYSPNRNEPLTGN 20
:|||||:|:|:|
Db 811 IDGYSEQVSEPIFSNS 826

RESULT 11
T16305
hypothetical protein F40F4.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16305
R:Wilson, R.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F40F4.
A:Reference number: Z18493
A:Accession: T16305
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2214 <WIL>
A:Cross-References: UNIPROT:Q20219; EMBL:U40420; NID:G1065513; PID:G1065514; PIDN:AAA814
C:Genetics:
A:Gene: CESP:F40F4.6
A:Introns: 57/3; 95/1; 302/3; 323/1; 380/1; 404/1; 468/1; 507/2; 547/3; 595/3; 1532/1; 1

Query Match
Best Local Similarity 43.3%; Score 45; DB 2; Length 2214;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 VLVDGSPNRNEPLNG 18
:|:|:|:|:|:|
Db 1950 VICNGRSPPTNOPATG 1965

RESULT 12
C81903
hypothetical protein NMA1343 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

```

```

C:Accession: C81903
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel]
; Holroyd, S.; Jagsis, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: C81903
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <PAR>
A:Cross-References: GB:AL162755; GB:AL157959; NID:G7379742; PIDN:CAB84591.1; PID:G738001;
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1343

Query Match
Best Local Similarity 42.3%; Score 44; DB 2; Length 186;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 DGYSPNRNEPLTGN 20
|||:|:|:|:|
Db 129 DGYTLPHHBAIAGNA 143

RESULT 13
D8118
conserved hypothetical protein / ankyrin-related protein NMB1133, NMB1171 [imported] - Ne
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: D8118; AB1114
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, B.K.; Hatt, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizzi, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: D8118
A:Molecule type: DNA
A:Residues: 1-253 <TET>
A:Cross-References: UNIPROT:Q9URZ6; GB:AE002462; GB:AE002098; NID:G7226363; PIDN:AAF41523
A:Experimental source: serogroup B, strain MC58
A:Accession: AB1114
A:Molecule type: DNA
A:Residues: 1-253 <TE3>
A:Cross-References: GB:AE002465; GB:AE002098; NID:G7226401; PIDN:AAF41556.1; PID:G7226401
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1133; NMB1171

Query Match
Best Local Similarity 42.3%; Score 44; DB 2; Length 253;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 DGYSPNRNEPLTGN 20
|||:|:|:|:|
Db 196 DGYTLPHHBAIAGNA 210

RESULT 14
AD2023
carboxyl-terminal proteinase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AD2023
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yabate, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2023
A:Status: preliminary

```

A:Molecule type: DNA
 A:Residues: 1-445 <KUR>
 A:Cross-references: UNIPROT:Q8YW78; GB:BA000019; PIDN:BA073437.1; PID:G17130828; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: a111738
 C:Superfamily: carboxyl-terminal processing protease

Query Match 42.3%; Score 44; DB 2; Length 445;
 Best Local Similarity 58.8%; Pred. No. 34;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SVLVDSPPNNEPLTG 18
 :|||||: :|||
 Db 312 AVLVDSNNSASSELTG 328

RESULT 15

T35818
 probable integral membrane export protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T35818
 R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, February 1999
 A:Reference number: Z21589
 A:Accession: T35818
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-748 <MUR>
 A:Cross-references: UNIPROT:Q9Z577; EMBL:AL035569; PIDN:CA037578.1; GSPDB:GN00070; SCOE
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOE09.14

Query Match 42.3%; Score 44; DB 2; Length 748;
 Best Local Similarity 53.3%; Pred. No. 62;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 VLVDSPPNNEPLT 17
 :|||:| :|||
 Db 408 LVLDSPPNNEPLT 422

Search completed: October 27, 2005, 18:57:23
 Job time : 13.3721 secs

(Order) Nymphaeaceae (Nymphaeaceae)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:34:49 ; Search time 58.1395 Seconds
(without alignments)
176.155 Million cell updates/sec

Title: US-10-612-090-14
Perfect score: 104
Sequence: 1 SSVLVGVGSPNNRNPITGNS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	1148	2 Q9H7S7	Q9H7S7 homo sapien
2	104	100.0	6995	2 Q6GRK2	Q6GRK2 homo sapien
3	104	100.0	22152	2 Q8WK17	Q8WK17 homo sapien
4	55	52.9	258	2 Q9D1H1	Q9D1H1 mus musculi
5	50	48.1	3519	1 Q9L56	Q9L56 STRAT
6	49	47.1	1666	1 CO3_CAVPO	CO3_CAVPO
7	48	46.2	1714	2 Q6I203	Q6I203 ostrinia nu
8	47	45.2	347	2 Q7VCG8	Q7VCG8 prochloroco
9	47	45.2	407	1 PPI_YEAST	PPI_YEAST
10	47	45.2	407	2 Q6Q5G6	Q6Q5G6 saccharomyc
11	47	45.2	867	2 Q9XBS2	Q9XBS2 zymomonas m
12	46	44.2	227	2 Q63WY6	Q63WY6 burkholderi
13	46	44.2	302	2 Q6CJ98	Q6CJ98 kluyveromyc
14	46	44.2	306	2 Q7NMG2	Q7NMG2 chromobacte
15	46	44.2	424	2 Q8AS18	Q8AS18 bacteroides
16	46	44.2	944	2 Q9GYG8	Q9GYG8 caenorhabdi
17	46	44.2	1018	1 M2R2_MOUSE	M2R2_MOUSE
18	46	44.2	1018	2 Q8BH85	Q8BH85 m musculus
19	46	44.2	1457	2 Q7R039	Q7R039 homologue
20	46	44.2	1723	2 Q9SWK9	Q9SWK9 giardia lam
21	45	43.3	107	1 UCRC_NEUCR	UCRC_NEUCR
22	45	43.3	167	1 HXD3_BRARE	HXD3_BRARE
23	45	43.3	242	2 Q89YJ4	Q89YJ4 bacteroides
24	45	43.3	271	2 Q8TV05	Q8TV05 methanopyru
25	45	43.3	324	2 Q7QH82	Q7QH82 anopheles g
26	45	43.3	325	2 Q7QH81	Q7QH81 anopheles g
27	45	43.3	449	2 Q9SR19	Q9SR19 arabidopsis
28	45	43.3	513	1 KCG3_HUMAN	KCG3_HUMAN
29	45	43.3	513	2 Q86T85	Q86T85 homo sapien
30	45	43.3	715	2 Q821Q0	Q821Q0 salmonella
31	45	43.3	1042	2 Q7REQ7	Q7REQ7 plasmodium

32	45	43.3	1046	2 Q9DXH5	Q9DXH5 human cytom
33	45	43.3	1048	1 P100_HCMVA	P08318 human cytom
34	45	43.3	1048	2 Q7M6Q1	Q7M6Q1 human cytom
35	45	43.3	1049	2 Q6RX16	Q6RX16 human cytom
36	45	43.3	1049	2 Q6SW99	Q6SW99 human cytom
37	45	43.3	1309	1 ST16_SCHPO	Q09743 schizosacch
38	45	43.3	2214	2 Q20219	Q20219 caenorhabdi
39	44.5	42.8	586	2 Q884Z9	Q884Z9 pseudomonas
40	44	42.3	155	2 Q8RMB3	Q8RMB3 uncultured
41	44	42.3	200	2 Q8S3A6	Q8S3A6 mycobacteri
42	44	42.3	215	1 PCP1_RALSO	Q8X69 ralbionia s
43	44	42.3	253	1 YB33_NEIMB	Q9J26 neisseria m
44	44	42.3	253	1 YD43_NEIMA	Q9J22 neisseria m
45	44	42.3	259	2 Q17188	Q17188 bombyx mori

ALIGNMENTS

RESULT 1
ID Q9H7S7 PRELIMINARY; PRT; 1148 AA.
AC Q9H7S7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14303.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Iehibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimura M., Watanabe S., Hirooka S., Chiba Y., Iehida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shichara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakabe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamagaki M., Watanabe K., Kumagai A., Itakura S., Fukumaki Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okumoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK024365; BAB14899.1; -.
DR HSSP; Q9D1H1; 11VZ.
DR InterPro: IPR000082; SEA.
DR Pfam; PF01390; SEA; 7.
DR PROSITE; PS50024; SEA; 3.
DR DR
SQ SEQUENCE 1148 AA; 127957 MW; 3861B0D5EFD8ABC CRC64;

```

Query Match          100.0%; Score 104; DB 2; Length 1148;
Best Local Similarity 100.0%; Pred. No. 4,3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SSVLDGYSPPNREPLTGNS 20
Db 1070 SSVLDGYSPPNREPLTGNS 1089

RESULT 2
O96RK2 PRELIMINARY; PRT; 6995 AA.
AC O96RK2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mucin 16 (Fragment).
GN Name=MUC16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200;
RY Yin B.W., Lloyd K.O.;
RT "Molecular cloning of the cal25 ovarian cancer antigen. Identification
as a new mucin, muc16."
RJ J. Biol. Chem. 276:27371-27375(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Lloyed K.O., Yin B.W.T.;
RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF361486; AA074120.3; -.
DR HSSP; Q9DIH1; 11VZ.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 20.
DR SMART; SMO0200; SEA; 10.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00024; SEA; 6.
FT NON TER
SO SEQUENCE 6995 AA; 744958 MW; 80C797BDBF33A2B CRC64;

Query Match          100.0%; Score 104; DB 2; Length 6995;
Best Local Similarity 100.0%; Pred. No. 3,7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SSVLDGYSPPNREPLTGNS 20
Db 6917 SSVLDGYSPPNREPLTGNS 6936

RESULT 3
O8WXI7 PRELIMINARY; PRT; 22152 AA.
AC O8WXI7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Ovarian cancer related tumor marker CA125.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
RY O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RA York L.;
RT "The CA 125 gene: an extracellular superstructure dominated by repeat
sequences."
RJ Tumour Biol. 22:348-366(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA O'Brien T.J., Underwood L.J., Beard J.B.;
RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF144442; AAL65133.2; -.
DR Gene; HGNC:15582; MOC16.
DR Pfam; PF01390; SEA; 51.
DR SMART; SMO0200; SEA; 23.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_2.
DR PROSITE; PS00024; SEA; 11.
SQ SEQUENCE 22152 AA; 235268 MW; B3E7BDF19997A440 CRC64;

Query Match          100.0%; Score 104; DB 2; Length 22152;
Best Local Similarity 100.0%; Pred. No. 1,4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SSVLDGYSPPNREPLTGNS 20
Db 22074 SSVLDGYSPPNREPLTGNS 22093

RESULT 4
O9DIH1 PRELIMINARY; PRT; 258 AA.
AC O9DIH1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:111000814 product:hypothetical SEA domain containing
DE protein, full insert sequence.
GN Name=111000814Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RL MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RJ Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RL MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RJ Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RJ Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RL MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RJ Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RL MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama T., Nishi K., Kitajima T., Tashiro H., Itoh M.,

```


RT "Amino acid sequence of guinea pig C3a anaphylatoxin.";
 RL Protein Seq. Data Anal. 1:473-478(1988).
 RN [3]
 RP SEQUENCE OF 993-1032.
 RX MEDLINE=83178889; PubMed=6938833;
 RA Thomas M.L., Tack B.F.;
 RT "Identification and alignment of a thiol ester site in the third component of guinea pig complement.";
 RT Biochemistry 22:942-947(1983).
 CC -1- FUNCTION: C3 plays a central role in the activation of the complement system. Its processing by C3 convertase is the central reaction in both classical and alternative complement pathways. After activation C3b can bind covalently, via its reactive thioester, to cell surface carbohydrates or immune aggregates.
 CC -1- FUNCTION: Derived from proteolytic degradation of complement C3, C3a anaphylatoxin is a mediator of local inflammatory process. It induces the contraction of smooth muscle, increases vascular permeability and causes histamine release from mast cells and basophilic leukocytes.
 CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg residues, forming two chains, beta and alpha, linked by a disulfide bond. C3 convertase activates C3 by cleaving the alpha chain, releasing C3a anaphylatoxin and generating C3b (beta chain + alpha' chain).
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib.ch).

 DR EMBL, M34054; AAA37038.1; -
 DR PIR, A37156; C3GP.
 DR HSSP, P01026; 10QF.
 DR InterPro: IPR002890; A2M_N.
 DR InterPro: IPR009048; AM_receptor_bind.
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR001840; Invasin_intimin.
 DR InterPro: IPR008964; Invasin_intimin.
 DR InterPro: IPR001599; MacroglobinA2.
 DR InterPro: IPR001134; Netrin_C.
 DR InterPro: IPR008930; Terp_Cyc_toroid.
 DR InterPro: IPR008993; TIMP_like.
 DR Pfam: PF00207; A2M; 1.
 DR Pfam: PF01835; A2M_N; 1.
 DR Pfam: PF01821; ANATO; 1.
 DR Pfam: PF01759; NTR; 1.
 DR PRINTS; PR00004; ANAPHYLATOXN.
 DR PRODOM; PD003264; Anaphylatoxin; 1.
 DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
 DR PROSITE; PS50189; NTR; 1.
 DR KW Complement alternate pathway; Complement pathway;
 KW Direct protein sequencing; Glycoprotein; Inflammatory response;
 KW Plasma; Signal; Thioester bond.
 FT SIGNAL 1 29
 FT CHAIN 30 1666 Complement C3.
 FT CHAIN 30 671 Complement C3 beta chain.
 FT CHAIN 676 1666 Complement C3 alpha chain.
 FT PEPTIDE 676 753 C3a anaphylatoxin.
 FT CHAIN 754 1666 Complement C3b alpha' chain.
 FT DOMAIN 698 723 Anaphylatoxin-like.
 FT DOMAIN 1522 1664 NTR.
 FT SITE 753 754 Cleavage (by C3 convertase).
 FT DISULFID 557 821 Interchain (By similarity).
 FT DISULFID 630 666 By similarity.
 FT DISULFID 698 725 By similarity.
 FT DISULFID 699 732 By similarity.

FT DISULFID 712 733 By similarity.
 FT DISULFID 878 1517 By similarity.
 FT DISULFID 1106 1163 By similarity.
 FT DISULFID 1363 1493 By similarity.
 FT DISULFID 1394 1462 By similarity.
 FT DISULFID 1510 1515 By similarity.
 FT DISULFID 1522 1593 By similarity.
 FT DISULFID 1540 1664 By similarity.
 FT DISULFID 1640 1649 By similarity.
 FT CROSSLINK 1015 1018 Isoglutamyl cysteine thioester (Cys-Gln).
 FT CARBOHYD 944 944 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 1620 1620 D->N (in Ref. 2).
 FT CONFLICT 731 731 Missing (in Ref. 3).
 FT CONFLICT 1013 1013 Q->E (in Ref. 2).
 FT CONFLICT 1018 1018 Missing (in Ref. 3).
 FT CONFLICT 1031 1031 Missing (in Ref. 3).
 SQ SEQUENCE 1666 AA; 186487 MW; 1C1F1219944AFD49 CRC64;
 Query Match 47.1%; Score 49; DB 1; Length 1666;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Db 3 VLVDGSPNRPNPLTNGS 20
 305 VLVDGSPNRPNPLTNGS 322
 RESULT 7
 ID 061203 PRELIMINARY; PRT; 1714 AA.
 AC 061203;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Cadherin A1.
 OS Oestrinia nubilalis (European corn borer).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyralidae;
 OC Pyralidae; Pyraustinae; Ostrinia.
 OX NCBI_Taxid=29057;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Coates B.S., Sumnerford D.V., Hellmich R.L.;
 RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Contains 11 cadherin domains.
 DR EMBL, AY612336; AAT31678.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; Cadherin; 3.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 10.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS50268; CADHERIN_2; 11.
 KW Calcium; Calcium-binding.
 SQ SEQUENCE 1714 AA; 191900 MW; 89E3CAE0614654DE CRC64;
 Query Match 46.2%; Score 48; DB 2; Length 1714;
 Best Local Similarity 55.6%; Pred. No. 1.6e+02;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Db 3 VLVDGSPNRPNPLTNGS 20
 1014 VLPEIYADRDDEPDTNS 1031
 RESULT 8
 ID 07VCG8 PRELIMINARY; PRT; 347 AA.
 AC 07VCG8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

01-MAR-2004 (TREMblrel. 26, last annotation update)
 DE Predicted ATPase.
 GN OrderedLocNames=Pro0772;
 OS Prochlorococcus marinus.
 OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcales;
 CC Prochlorococcus.
 NCBI_TaxID=1219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SARG / CCMF 1375 / SS120;
 RX MEDLINE=22810154; PubMed=12917466; DOI=10.1073/pnas.1733211100;
 RA Dufresne A., Salanoubat M., Partenecky F., Artiguenave F., Axmann I.M.,
 Barde V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
 Makarova K.S., Ostrowski M., Oztra S., Robert C., Rogozin I.B.,
 Scanlan D.J., Tandeau de Marsac N., Weissbach J., Wincker P.,
 Wolf Y.I., Hesse W.R.;
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
 a nearly minimal oxyphototrophic genome."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
 DR EMBL; AE017163; AAP9916.1; -.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR001307; Thiosulfatrans.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS00380; RHODANSE_1; UNKNOWN_1.
 DR PROSITE; PS00206; RHODANSE_3; 1.
 DR Complete proteome.
 SQ SEQUENCE 347 AA; 39187 MW; 388F2B1A2BE4956 CRC64;

Query Match 45.2%; Score 47; DB 2; Length 347;
 Best Local Similarity 69.2%; Pred. No. 35;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LVGVSFNRNEPL 16
 DB 258 LVDFVSFNRNEPL 270

RESULT 9
 RPI1_YEAST STANDARD; PRT; 407 AA.
 AC P23250;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Negative RAS protein regulator protein.
 GN Name=RPI1; OrderedLocNames=YIL119C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91304380; PubMed=1649384;
 RA Kim J.-H., Powers S.;
 RT "Overexpression of RPI1, a novel inhibitor of the yeast Ras-cyclic AMP
 pathway, down-regulates normal but not mutationally activated ras
 function."
 RT Mol. Cell. Biol. 11:3894-3904(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6286C / AB972;
 RX MEDLINE=97313266; PubMed=9169870;
 RA Churcher C.M., Bowden S., Badcock K., Bankier A., Brown D.,
 Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 Harris D.E., Horell T., Hunt S., Jagels K., Jones M., Lye G.,
 Moutre S., O'Dell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX."
 RL Nature 387:84-87(1997).
 CC -!- FUNCTION: Negative regulator of the Ras-cyclic AMP pathway.
 CC Negatively regulate the activity of normal but not mutationally
 CC activated Ras proteins. The down-regulatory effect of RPI1
 CC requires the presence of one of the two Ras GTPase activators,

IRAI and IRA2.
 CC SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 Myb-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 246833; CAA6873.1; -
 CC EMBL; M63178; AAA35001.1; -
 CC EMBL; 542682; AAA39843.1; -
 CC PIR; S49890; S49890.
 CC Germonline; 139654; -
 CC SGD; S000001381; RPI1.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IMP.
 CC GO; GO:0007265; P:RAS protein signal transduction; IMP.
 CC InterPro; IPR001005; Myb_DNA binding.
 CC PROSITE; PS00037; MYB_1; FALSE_NEG.
 CC PROSITE; PS00334; MYB_2; FALSE_NEG.
 CC DR PROSITE; PS50090; MYB_3; 1.
 CC KM DNA-binding; Nuclear protein.
 FT DNA BIND 90 158
 FT MYB.
 FT N-> K (in Ref. 1).
 FT CONFLICT 37 37
 FT CONFLICT 71 73
 FT SNS-> P (in Ref. 1).
 FT CONFLICT 91 91
 FT A-> V (in Ref. 1).
 FT CONFLICT 170 174
 FT Missing (in Ref. 1).
 FT CONFLICT 240 240
 FT C-> S (in Ref. 1).
 FT CONFLICT 242 242
 FT N-> S (in Ref. 1).
 FT CONFLICT 252 252
 FT N-> NNNSNN (in Ref. 1).
 SQ SEQUENCE 407 AA; 46623 MW; C6A18DCD04F58701 CRC64;

Query Match 45.2%; Score 47; DB 1; Length 407;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 YSPNRNEPLTGN 19
 DB 26 YSPNLTPTTGN 37

RESULT 10
 Q605G6 PRELIMINARY; PRT; 407 AA.
 AC Q605G6;
 DT 05-JUL-2004 (TREMblrel. 27, Created)
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
 DE YIL119C.
 GN Name=RPI1;
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marsischky G., Rolfs A., Richardson A., Kane M., Bagui M., Taycher E.,
 Hu Y., Vanberg J., Weger J., Kramer J., Moreira D., Kelley F.,
 Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
 Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
 Labber J.;
 RT Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY558088; AAS56414.1; -
 DR SGD; S000001381; RPI1.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR001005; Myb_DNA binding.
 DR PROSITE; PS50090; MYB_3; 1.
 SQ SEQUENCE 407 AA; 46633 MW; 448D6C302815B6FD CRC64;

```

Query Match          45.2%; Score 47; DB 2; Length 407;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 YSPNREPLTGN 19
   ||||| :|||
Db 26 YSPNINPTITSN 37

RESULT 11
OQXBS2
ID OQXBS2; PRELIMINARY; PRT; 867 AA.
AC OQXBS2;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Membrane alanyl aminopeptidase.
GN Name=pepN;
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZM4;
RA Um H.W., Kang H.S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF157493; AAD42403.1; -.
DR MEROPS; M01.005; -.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001930; Peptidase_M1.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PRINTS; PR00756; ALADIPRASE.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR Aminopeptidase.
SQ SEQUENCE 867 AA; 97500 MW; BDD94757AC90BCC CRC64;

Query Match          45.2%; Score 47; DB 2; Length 867;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 VLVDSYSPNRNPLT 17
   ||||| :|||
Db 713 VLVNGFSPERKALT 727

RESULT 12
OQ3WY6
ID OQ3WY6; PRELIMINARY; PRT; 227 AA.
AC OQ3WY6;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BPSL0753;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Tiltball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Baeson N., Beesham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Felwell T., Fraser A., Hance Z., Hauser C., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,

```

```

RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songstivalai S., Stevens K., Tumapa S., Vesaratchaveat M.,
RA Whitehead S., Yeats C., Barrett B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH34746.1; -.
KW Hypothetical protein.
SQ SEQUENCE 227 AA; 24840 MW; D98FA4F04FECB059 CRC64;

```

```

Query Match          44.2%; Score 46; DB 2; Length 227;
Best Local Similarity 61.5%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 VDGYSPPNRNPLT 17
   ||||| :|||
Db 121 IDGYSKNSESPNT 133

```

RESULT 13

```

OQ6CJ98
ID OQ6CJ98; PRELIMINARY; PRT; 302 AA.
AC OQ6CJ98;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Similar to egd|S0005690 Saccharomyces cerevisiae YOR164C.
GN ORFNames=KLAD0F202959;
OS Kluyveromyces lactic NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykaeten C.,
RA Boerame A., Boyer J., Catolico L., Confantoleri F., de Darvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bojotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Galliardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382126; CAG38699.1; -.
DR InterPro; IPR007317; DUF410.
DR Pfam; PF04190; DUF410; 1.
SQ SEQUENCE 302 AA; 34920 MW; 0C295541A6165E00 CRC64;

```

```

Query Match          44.2%; Score 46; DB 2; Length 302;
Best Local Similarity 56.2%; Pred. No. 44;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 3 VLVDSYSPNRNPLT 18
   ||||| :|||
Db 95 VLVDSQEPNLDVVTG 110

```

RESULT 14

Q7NMKG2

```
ID 07NMG2 PRELIMINARY; PRT; 306 AA.
AC 07NMG2;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Chemotaxis motB protein.
GN Name=mocB2; OrderedLocustNames=CV2027;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RC MEDLINE=22882880; Pubmed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimarães C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Aratipe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Batuas L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burily H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
RA Crezymbek-Tasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Gonçalves P.R., Grangeiro T.B.,
RA Gracatapaglia D., Grissard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leol L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manlio G.P., Maranhao A.O., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunares H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Umenyi T.,
RA Vettore A., Waesem R., Zana A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DB EMBL: AE016917; AA059699.1;
DB GO: GO:0009779; Counter membrane (sensu Gram-negative Bacteria); IEA.
DB InterPro: IPR006665; OmpA/MocB.
DB Pfam: PF00691; OmpA; 1.
DB ProDom: PD000930; OmpA/MocB; 1.
DB Complete proteome.
SQ SEQUENCE 306 AA; 33083 MW; D2F4EDBDP832A18 CRC64;
```

```
Query Match 44.2%; Score 46; DB 2; Length 306;
Best Local Similarity 44.4%; Pred. No. 45;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 SSVLVGCGSPNRPNEPLTG 18
DB 273 SSVLVGCGSPNRPNEPLTG 290
```

```
RESULT 15
Q8A5L8 PRELIMINARY; PRT; 424 AA.
AC Q8A5L8;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BT2220;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RC MEDLINE=22550858; Pubmed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjurell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DB EMBL: AE016935; AA077327.1;
DB Complete proteome; Hypothetical protein.
SQ SEQUENCE 424 AA; 47549 MW; 3C87947B0184AC3E CRC64;
```

```
Query Match 44.2%; Score 46; DB 2; Length 424;
Best Local Similarity 40.0%; Pred. No. 66;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 SSVLVGCGSPNRPNEPLTGS 20
DB 358 SSVLVGCGSPNRPNEPLTGS 377
```

```
Search completed: October 27, 2005, 18:55:07
Job time : 63.1395 secs
```

Two Page Book (1990)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:31:33 ; Search time 69.2093 Seconds
(without alignments)
111.765 Million cell updates/sec

Title: US-10-612-090-15
Perfect score: 105
Sequence: 1 TNYQRKRNIEDALNQLFRN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	20	ADJ66742	Adj66742 Human Muc
2	105	100.0	108	ADJ66729	Adj66729 Human Muc
3	105	100.0	108	ADJ66740	Adj66740 Human Muc
4	105	100.0	178	ABP30979	Abp30979 Predicted
5	105	100.0	178	ADA08642	Ada08642 Human O77
6	105	100.0	178	Adf08985	Adf08985 Secreted
7	105	100.0	233	ABP30978	Abp30978 Extracellular
8	105	100.0	233	ADA08641	Ada08641 Human O77
9	105	100.0	233	Adf08984	Adf08984 Secreted
10	105	100.0	284	ABU54859	Abu54859 Human CAL
11	105	100.0	318	ABP31025	Abp31025 Amino aci
12	105	100.0	318	ADA08747	Ada08747 Human O77
13	105	100.0	318	Adf09090	Adf09090 Secreted
14	105	100.0	367	AAE12631	Aae12631 Human ova
15	105	100.0	438	AAE12555	Aae12555 Human ova
16	105	100.0	438	AAE12623	Aae12623 Human gen
17	105	100.0	438	ABP30973	Abp30973 Partial p
18	105	100.0	438	ABP30899	Abp30899 0772P c10
19	105	100.0	438	ADA08636	Ada08636 Human O77
20	105	100.0	438	ADA08543	Ada08543 Human ova
21	105	100.0	438	Adf08886	Adf08886 Secreted
22	105	100.0	438	Adf08979	Adf08979 Secreted
23	105	100.0	438	ADG46173	Adg46173 Human ova
24	105	100.0	439	ABU54709	Abu54709 Human CAL
25	105	100.0	526	AAE12634	Aae12634 Human gen

26	105	100.0	748	8	ADS94302	Ads94302 CA 125/O7
27	105	100.0	809	3	ADS94303	Ads94303 CA 125/O7
28	105	100.0	833	3	AAE12554	Aae12554 Human ova
29	105	100.0	833	5	ABP30898	Abp30898 0772P c10
30	105	100.0	833	7	ADA08542	Ada08542 Human ova
31	105	100.0	833	7	Adf08885	Adf08885 Secreted
32	105	100.0	833	7	ADG46172	Adg46172 Human ova
33	105	100.0	914	3	AAE12552	Aae12552 Human ova
34	105	100.0	914	4	ABP99203	Abp99203 Human ova
35	105	100.0	914	5	ABP30968	Abp30968 Hypothetel
36	105	100.0	914	5	ABP30896	Abp30896 0772P pro
37	105	100.0	914	7	ADA08631	Ada08631 Human O77
38	105	100.0	914	7	ADA08465	Ada08465 Human ova
39	105	100.0	914	7	Adf08974	Adf08974 Secreted
40	105	100.0	914	7	Adf08808	Adf08808 Secreted
41	105	100.0	914	7	ADG46095	Adg46095 Human ova
42	105	100.0	914	8	ADN40451	Adn40451 Human bre
43	105	100.0	1148	4	AAE12583	Aae12583 Human pro
44	105	100.0	1148	4	ABP50283	Abp50283 HOST-1 ov
45	105	100.0	1148	5	ABG96380	Abg96380 Human ova

ALIGNMENTS

RESULT 1
ADJ66742
ID ADJ66742 standard; peptide; 20 AA.
XX
AC ADJ66742;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human Muc1/Muc16 protein-related peptide SegID15.
XX
KW monoclonal antibody; epitope; non-shed extracellular portion;
KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
XX breast cancer; ovarian cancer.
XX
OS Homo sapiens.
XX
PN WO2004005470-A2.
XX
PD 15-JAN-2004.
XX
PF 03-JUL-2003; 2003WO-US020907.
XX
PR 03-JUL-2002; 2002US-0393094P.
XX
PA (IMMU-) IMMUNOGEN INC.
XX
PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
PI Water CA;
XX
DR WPI; 2004-091350/09.
XX
PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
PT treating or monitoring malignancies, such as breast or ovarian cancer.
XX
PS Claim 14; SEQ ID NO 15; 113pp; English.
XX
CC This invention relates to a novel isolated monoclonal antibody that
CC specifically binds to an epitope of a non-shed extracellular portion of a
CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
CC useful for the development of compounds with a cytostatic or for gene
CC therapy. The composition and methods are useful in diagnosing, treating
CC or monitoring malignancies, such as breast or ovarian cancer. The present
CC sequence is that of a peptide which was used for raising antibodies to
CC the extracellular, non-shed region of Muc16 and which is claimed in the
CC specification.
XX
XX Sequence 20 AA;
SQ

Query Match 100.0%; Score 105; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.6e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TNYORKNRIEDALNQLFRN 20
 |||||
 1 TNYORKNRIEDALNQLFRN 20

RESULT 2

ADJ66729
 ID ADJ66729 standard; protein; 108 AA.

AC ADJ66729;

DT 06-MAY-2004 (first entry)

DE Human Muc16 epitope amino acid sequence.

KM monoclonal antibody; epitope; non-shed extracellular portion;
 KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KM breast cancer; ovarian cancer.

OS Homo sapiens.

PN WO2004005470-A2.

PD 15-JAN-2004.

PF 03-JUL-2003; 2003WO-US020907.

PR 03-JUL-2002; 2002US-0393094P.

PA (IMMU-) IMMUNOGEN INC.

PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
 PI Vater CA;

DR WPI; 2004-091350/09.

PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
 PT treating or monitoring malignancies, such as breast or ovarian cancer.

PS Claim 13; SEQ ID NO 2; 113pp; English.

CC This invention relates to a novel isolated monoclonal antibody that
 CC specifically binds to an epitope of a non-shed extracellular portion of a
 CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
 CC useful for the development of compounds with a cytostatic or for gene
 CC therapy. The composition and methods are useful in diagnosing, treating
 CC or monitoring malignancies, such as breast or ovarian cancer. The present
 CC sequence is that of a Muc epitope of the invention.

SQ Sequence 108 AA;

Query Match 100.0%; Score 105; DB 8; Length 108;
 Best Local Similarity 100.0%; Pred. No. 6.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TNYORKNRIEDALNQLFRN 20
 |||||
 1 TNYORKNRIEDALNQLFRN 20

RESULT 3

ADJ66740
 ID ADJ66740 standard; protein; 108 AA.

AC ADJ66740;

DT 06-MAY-2004 (first entry)

DE Human Muc16 GST fusion protein amino acid sequence SeqID13.

XX monoclonal antibody; epitope; non-shed extracellular portion;
 KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KW breast cancer; ovarian cancer.

OS Homo sapiens.

PN WO2004005470-A2.

PD 15-JAN-2004.

PF 03-JUL-2003; 2003WO-US020907.

PR 03-JUL-2002; 2002US-0393094P.

PA (IMMU-) IMMUNOGEN INC.

PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
 PI Vater CA;

DR WPI; 2004-091350/09.

PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
 PT treating or monitoring malignancies, such as breast or ovarian cancer.

PS Disclosure; SEQ ID NO 13; 113pp; English.

CC This invention relates to a novel isolated monoclonal antibody that
 CC specifically binds to an epitope of a non-shed extracellular portion of a
 CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
 CC useful for the development of compounds with a cytostatic or for gene
 CC therapy. The composition and methods are useful in diagnosing, treating
 CC or monitoring malignancies, such as breast or ovarian cancer. The present
 CC sequence is that of the human Muc16 GST fusion protein which was used in
 CC the exemplification of the invention.

SQ Sequence 108 AA;

Query Match 100.0%; Score 105; DB 8; Length 108;
 Best Local Similarity 100.0%; Pred. No. 6.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TNYORKNRIEDALNQLFRN 20
 |||||
 1 TNYORKNRIEDALNQLFRN 20

RESULT 4

ABP30979
 ID ABP30979 standard; protein; 178 AA.

AC ABP30979;

DT 02-JUL-2002 (first entry)

DE Predicted extracellular domain of 0772P.

KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.

OS Homo sapiens.

PN WO200206317-A2.

PD 24-JAN-2002.

PF 17-JUL-2001; 2001WO-US022635.

PR 17-JUL-2000; 2000US-00617747.

PR 10-AUG-2000; 2000US-00636801.

PR 20-SEP-2000; 2000US-00667857.

PR 04-APR-2001; 2001US-00827271.

PR 18-JUN-2001; 2001US-00884441.

PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2002-164781/21.
DR
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
XX Example 13; Page 361-362; 408pp; English.
XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations, and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX
SQ Sequence 178 AA:

Query Match 100.0%; Score 105; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYQRKNRNIEDALNQLFRN 20
DB 70 TNYQRKNRNIEDALNQLFRN 89

RESULT 5

ADA08642
ID ADA08642 standard; protein; 178 AA.

AC ADA08642;

XX 06-NOV-2003 (first entry)

XX Human O772P partial protein #12.

XX human; gene therapy; ovarian cancer; cancer.

XX Homo sapiens.

XX US2003091580-A1.

XX 15-MAY-2003.

XX 17-JUL-2001; 2001US-00907969.

XX 18-JUN-2001; 2001US-00884441.

XX (MITC/) MITCHAM J L.

XX (KING/) KING G E.

XX (ALGA/) ALGATE P A.

XX (FLIN/) FLING S P.

XX (RETT/) RETTER M W.

XX (FANG/) FANGER G R.

XX (REED/) REED S G.

XX (VEDV/) VEDVICK T S.

XX (CART/) CARTER D.

XX (HILL/) HILL P.

XX (ALBO/) ALBONE E.

XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.

XX New isolated O772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian

PT cancer.
XX
XX Disclosure; Page 28; 371pp; English.
XX
XX The invention relates to an isolated O772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen.
XX
SQ Sequence 178 AA:

Query Match 100.0%; Score 105; DB 7; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYQRKNRNIEDALNQLFRN 20
DB 70 TNYQRKNRNIEDALNQLFRN 89

RESULT 6
ADF08985
ID ADF08985 standard; protein; 178 AA.

AC ADF08985;

XX 12-FEB-2004 (first entry)

XX Secreted ovarian carcinoma antigen seqid 489.

XX gene therapy; protein therapy; vaccine; antibody inhibition;

XX breast cancer; restorative therapy; diagnostic agent; immunoassay;

XX secreted ovarian carcinoma antigen.

XX Homo sapiens.

XX US2003124140-A1.

XX 03-JUL-2003.

XX 17-JUL-2002; 2002US-00198053.

XX 17-DEC-1998; 98US-00215681.

XX 17-DEC-1998; 98US-00216003.

XX 23-JUN-1999; 99US-00338933.

XX 24-SEP-1999; 99US-00404879.

XX 17-JUL-2000; 2000US-00617747.

XX 10-AUG-2000; 2000US-00636801.

XX 20-SEP-2000; 2000US-00667857.

XX 04-APR-2001; 2001US-00827271.

XX 18-JUN-2001; 2001US-00884441.

XX 17-JUL-2001; 2001US-00907969.

XX (CORI-) CORIXA CORP.

XX Bangur CS, Retter MW, Fanger GR, Hill P;

XX WPI; 2003-697152/82.

XX Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
XX Example 13; SEQ ID NO 489; 399pp; English.

XX The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders

CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of (II) by expressing
 CC inactive proteins or to supplement the patient's own production of (II).
 CC Additionally, (I) may be used to produce (II), by inserting (I) into a
 CC host cell and culturing the cell to express the protein (II). (I) And its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. The host cell may also be used as antigens in the production of
 CC antibodies against (II) and in assays to identify modulators of (II)'s
 CC expression and activity. The anti-(II) antibodies, agonists and
 CC antagonists may be used to regulate expression and activity and as
 CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
 CC immunosassay). This sequence represents a secreted ovarian carcinoma
 CC antigen.

CC Sequence 178 AA:

Query Match 100.0%; Score 105; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYQRKNRINEDALNQLFRN 20
 |||
 Db 70 TNYQRKNRINEDALNQLFRN 89

RESULT 7
 ABP30978
 ID ABP30978 standard; protein; 233 AA.
 XX ABP30978;
 DT 02-JUL-2002 (first entry)
 XX Extracellular and transmembrane regions of 0772P.
 DE Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX Homo sapiens.
 OS WO200206317-A2.
 XX 24-JAN-2002.
 PD 17-JUL-2001; 2001WO-US022635.
 PF 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Mitcham JI, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 XX WPI; 2002-164781/21.
 DR
 XX
 XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 PS Example 13; Page 361; 408pp; English.
 CC This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents protein
 CC related to the invention

XX Sequence 233 AA;
 SQ Query Match 100.0%; Score 105; DB 5; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYQRKNRINEDALNQLFRN 20
 |||
 Db 70 TNYQRKNRINEDALNQLFRN 89

RESULT 8
 ADA08641
 ID ADA08641 standard; protein; 233 AA.
 XX ADA08641;
 AC
 XX
 XX 06-NOV-2003 (first entry)
 DT
 XX Human 0772P partial protein #11.
 DE
 XX human; gene therapy; ovarian cancer; cancer.
 KM
 XX Homo sapiens.
 OS
 XX US2003091580-A1.
 PN
 XX 15-MAY-2003.
 PD
 XX 17-JUL-2001; 2001US-00907969.
 PF
 XX 18-JUN-2001; 2001US-00884441.
 PR
 XX (MITC/) MITCHAM J L.
 PA (KING/) KING G E.
 PA (ALGA/) ALGATE P A.
 PA (FLIN/) FLING S P.
 PA (RETT/) RETTER M W.
 PA (FANG/) FANGER G R.
 PA (REED/) REED S G.
 PA (VEDV/) VEDVICK T S.
 PA (CART/) CARTER D.
 PA (HILL/) HILL P.
 PA (ALBO/) ALBONE E.
 XX
 PI Mitcham JI, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 XX WPI; 2003-532352/50.
 DR
 XX
 XX New isolated 0772P polypeptides and polynucleotides, useful in gene
 PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
 PT cancer.
 XX
 PS Example 13; Page 28; 371pp; English.
 CC The invention relates to an isolated 0772P polypeptide, which has the
 CC structure fully defined in the specification. The composition containing
 CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
 CC or antigen presenting cells are useful for stimulating an immune response
 CC and treating ovarian cancer. Detecting the presence of the
 CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
 CC carcinoma cDNAs and protein cDNAs were identified using microarray
 CC technology. The present sequence represents a human ovarian carcinoma
 CC antigen.
 CC
 XX Sequence 233 AA;
 SQ Query Match 100.0%; Score 105; DB 7; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Fri Oct 28 08:46:45 2005

QY 1 TNYQRKRNIEDALNQLFRN 20
 |||||
 DB 70 TNYQRKRNIEDALNQLFRN 89

RESULT 9
 ID ADF08984 standard; peptide; 233 AA.
 AC ADF08984;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Secreted ovarian carcinoma antigen seqid 488.
 XX
 KW Gene therapy; protein therapy; vaccine; antibody inhibition;
 KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
 KW secreted ovarian carcinoma antigen.
 XX
 OS Homo sapiens.
 XX
 PN US2003124140-A1.
 PD 03-JUL-2003.
 PF 17-JUL-2002; 2002US-00198053.
 XX
 PR 17-DEC-1998; 98US-00215681.
 PR 17-DEC-1998; 98US-00216003.
 PR 23-JUN-1999; 99US-00338933.
 PR 24-SEP-1999; 99US-00404879.
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00657857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 PR 17-JUL-2001; 2001US-00907969.
 XX
 PA (CORI-) CORIXA CORP.
 PI Bangur CS, Retter MW, Fanger GR, Hill P,
 DR WPI; 2003-897152/82.
 XX
 PT Oncogenic nucleic acids useful for the prevention, diagnosis and
 PT treatment of breast cancer.
 XX
 PS Example 13; SEQ ID NO 488; 399pp; English.
 XX
 CC The invention describes nucleic acids (I) and the polypeptides (II) they
 CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
 CC treating diseases related to their aberrant expression i.e. breast
 CC cancers. For example, (I) and (II) may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of (II) by expressing
 CC inactive proteins or to supplement the patient's own production of (II).
 CC Additionally, (II) may be used to produce (II), by inserting (I) into a
 CC host cell and culturing the cell to express the protein (II). (I) And its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. The host cell may also be used as antigens in the production of
 CC antibodies against (II) and in assays to identify modulators of (II)'s
 CC expression and activity. The anti-(II) antibodies, agonists and
 CC antagonists may be used to regulate expression and activity and as
 CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
 CC immunoassay). This sequence represents a secreted ovarian carcinoma
 CC antigen.
 XX
 SO Sequence 233 AA;
 Query Match 100.0%; Score 105; DB 7; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TNYQRKRNIEDALNQLFRN 20
 |||||
 DB 70 TNYQRKRNIEDALNQLFRN 89

RESULT 10
 ID ABUS4859 standard; protein; 284 AA.
 AC ABUS4859;
 XX
 DT 12-MAR-2003 (first entry)
 XX
 DE Human CA125 carboxy terminal domain.
 XX
 KW Human; CA125; protein repeat; chromosome 19q 13.2; amino terminal domain;
 KW amino terminal extension; carboxy terminal domain; vaccine; cancer;
 KW ovarian cancer; carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200283866-A2.
 PD 24-OCT-2002.
 PF 12-APR-2002; 2002WO-US011734.
 XX
 PR 17-APR-2001; 2001US-0284175P.
 PR 19-JUN-2001; 2001US-0299380P.
 PR 27-SEP-2001; 2001US-00965738.
 PR 21-DEC-2001; 2001US-0345180P.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 PI O'brien T, Beard J, Underwood L;
 DR WPI; 2003-093013/08.
 XX
 PT New CA125 molecules, useful as a gold standard for detecting and
 PT monitoring the presence of CA125 antigen which can be used for
 PT diagnosing, monitoring or treating patients with cancer or for developing
 PT vaccine against cancer.
 XX
 PS Claim 1; Fig 9b; 694pp; English.
 XX
 CC The invention relates to a CA125 protein comprising: (a) an extracellular
 CC amino terminal domain; (b) an amino terminal extension; (c) a multiple
 CC repeat domain; and (d) a carboxy terminal domain. The extracellular amino
 CC terminal domain comprises 5 genomic exons, the amino terminal extension
 CC comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and
 CC the carboxy terminal domain comprises a transmembrane anchor with a short
 CC cytoplasmic domain, and further comprises 9 genomic exons. The gene for
 CC CA125 is located on human chromosome 19q 13.2. Also included are isolated
 CC CA125 repeat domains, nucleic acids (including variants, homologues and
 CC degenerate versions) encoding CA125 proteins or repeat units, a vector
 CC comprising the nucleic acid, a cultured cell comprising the vector, a
 CC method of expressing CA125 antigen in a cell, the amino acid sequences of
 CC the CA125 repeat units (or their variants, fragments or sequences 50%
 CC identical to them), a purified antibody that selectively binds to an
 CC epitope in the receptor-binding domain of CA125 protein, a diagnostic for
 CC detecting and monitoring the presence of CA125 antigen (comprising
 CC recombinant CA125 having at least one repeat unit of the CA125 repeat
 CC domain including epitope binding sites), a therapeutic vaccine to treat
 CC mammals with elevated CA125 antigen levels or at risk of developing a
 CC disease or disease recurrence associated with elevated CA125 antigen
 CC levels (comprising recombinant CA125 repeat domains including epitope
 CC binding sites) and an antisense oligonucleotide that inhibits the
 CC expression of CA125. The CA125 molecule, particularly the multiple repeat
 CC domains are useful as a gold standard for detecting and monitoring the
 CC presence of CA125 antigen, which can be used for diagnosing, monitoring
 CC or treating patients with ovarian cancer and other carcinomas where CA125

CC is expressed. The molecules are also useful for developing a vaccine
 CC against cancer. The present sequence is a CA125 repeat protein
 XX
 SQ Sequence 284 AA;

Query Match 100.0%; Score 105; DB 6; Length 284;
 Best Local Similarity 100.0%; Pred. No. 2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TNYQRKNIEDALNQLFRN 20
 |||||
 DB 121 TNYQRKNIEDALNQLFRN 140

RESULT 11
 ABP31025
 ID ABP31025 standard; protein; 318 AA.
 XX
 AC ABP31025;

DT 02-JUL-2002 (first entry)

DE Amino acid sequence of the 3' constant region of 0772P.

KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.

XX Homo sapiens.

OS WO200206317-A2.

PN 24-JAN-2002.

PF 17-JUL-2001; 2001WO-US022635.

PR 17-JUL-2000; 2000US-00617747.

PR 10-AUG-2000; 2000US-00636801.

PR 20-SEP-2000; 2000US-00657857.

PR 04-APR-2001; 2001US-00827271.

PR 18-JUN-2001; 2001US-00884441.

XX (CORI-) CORIXA CORP.

PI Mitcham JI, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;

PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;

DR WPI; 2002-164781/21.

PS Claim 3; Page 399-400; 408pp; English.

CC This invention relates to polypeptides comprising an immunogenic portion

CC of an ovarian carcinoma protein which acts as an immunostimulant and is

CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations and antigen presenting cells that express

CC the polypeptides are useful for stimulating an immune response in a

CC patient and treating ovarian cancer. This sequence represents protein

CC related to the invention

XX Sequence 318 AA;

Query Match 100.0%; Score 105; DB 5; Length 318;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TNYQRKNIEDALNQLFRN 20
 |||||
 DB 155 TNYQRKNIEDALNQLFRN 174

RESULT 12

ADA08747
 ID ADA08747 standard; protein; 318 AA.
 XX
 AC ADA08747;

DT 06-NOV-2003 (first entry)

DE Human 0772P constant region consensus sequence #1.

XX human; gene therapy; ovarian cancer; cancer.

OS Homo sapiens.

PN US2003091580-A1.

PD 15-MAY-2003.

PF 17-JUL-2001; 2001US-00907969.

PR 18-JUN-2001; 2001US-00884441.

PA (MITC/) MITCHAM J L.

PA (KING/) KING G E.

PA (ALGA/) ALGATE P A.

PA (FLIN/) FLING S P.

PA (RETT/) RETTER M W.

PA (FANG/) FANGER G R.

PA (REED/) REED S G.

PA (VEDV/) VEDVICK T S.

PA (CART/) CARTER D.

PA (HILL/) HILL P.

PA (ALBO/) ALBONE E.

XX Mitcham JI, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;

PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;

DR WPI; 2003-53252/50.

XX New isolated 0772P polypeptides and polynucleotides, useful in gene

PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian

PT cancer.

XX Claim 3; SEQ ID NO 594; 371pp; English.

PS The invention relates to an isolated 0772P polypeptide, which has the

CC structure fully defined in the specification. The composition containing

CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells

CC or antigen presenting cells are useful for stimulating an immune response

CC and treating ovarian cancer. Detecting the presence of the

CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian

CC carcinoma cDNAs and protein cDNAs were identified using microarray

CC technology. The present sequence represents a human ovarian carcinoma

CC antigen.

XX Sequence 318 AA;

Query Match 100.0%; Score 105; DB 7; Length 318;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TNYQRKNIEDALNQLFRN 20
 |||||
 DB 155 TNYQRKNIEDALNQLFRN 174

RESULT 13
 ADF09090
 ID ADF09090 standard; protein; 318 AA.
 XX
 AC ADF09090;

DT 12-FEB-2004 (first entry)

XX 03-JAN-2002 (first entry)
XX
XX
DE Human gene 1 encoded secreted protein fragment, SEQ ID NO: 27.
XX
XX Human: ovarian cancer antigen; proliferative disorder; cancer; tumour;
KW respiratory system disorder; asthma; haematopoietic disorder; skin aging;
KW immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
KW rheumatoid arthritis; inflammation; neurological disorder; septic shock;
KW Alzheimer's disease; Parkinson's disease; diabetes; angiogenic disorder;
KW atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
KW epithelial cell proliferation; transplantation; chemotaxis; infection;
KW food additive; wound healing; endocrine disorder; kidney disorder;
KW gene therapy; cytostatic.
XX
XX Homo sapiens.
XX
XX W0200170804-A1.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001WO-US008585.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX 23-AUG-2000; 2000US-0227009P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI, 2001-639119/73.
XX
XX
XX Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
PT metastases.
XX
XX
XX Disclosure; Page 9; 427pp; English.
XX
XX AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
XX ovarian cancer associated protein (collectively known as ovarian cancer
XX antigens) genes; and AAE12623-AA12629 represent the proteins they
XX encode. AAE12630-AAE12638 represent human ovarian cancer antigenic
XX fragments or variants. Ovarian cancer antigens and their corresponding
XX DNAs are used in the prevention, diagnosis and treatment of diseases
XX associated with their inappropriate expression. These disorders include
XX proliferative disorders, cancer, tumours, respiratory system disorders,
XX asthma, haematopoietic disorders, diseases of the immune system, AIDS,
XX skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
XX inflammation, allergies, neurological disorders (e.g., Alzheimer's
XX disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, endocrine disorders and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, to identify their cognate ligands or binding
XX partners, in chemotaxis and can be used as a food additive. Antibodies
XX specific for a protein of the invention can be used in alleviating
XX symptoms associated with the disorders mentioned above and in diagnostic
XX immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
XX the invention is used in gene therapy. The present sequence represents a
XX human ovarian cancer antigenic fragment of the invention
XX
XX
XX Sequence 367 AA:
SO
Query Match 100.0%; Score 105; DB 4; Length 367;
Best Local Similarity 100.0%; Pred. No. 2, 7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNYQNRKNTEDALNQLFRN 20
db [|||||]|||||
275 TNYQNRKNTEDALNQLFRN 294

```

RESULT 15
AAB12555
ID AAB12555 standard; protein; 438 AA.
XX
AC AAB12555;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen O772P protein SEQ ID NO:390.
XX
KM Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
XX tumour antigen; identification; cytostatic; gene therapy; vaccine.
OS Homo sapiens.
XX
PN MO20036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US030270.
XX
PR 17-DEC-1998; 98US-00215661.
XX PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
DR N-PSDB; AAA70076.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer.
XX
PS Example 2; Page 203-204; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC cancer. AA69691 to AA70077 and AAB12552 to AAB12557 represent human
CC ovarian carcinoma polynucleotides and proteins used in the
CC exemplification of the present invention
XX
SQ Sequence 438 AA;

```

Query Match 100.0%; Score 105; DB 3; Length 438;
 Best Local Similarity 100.0%; Pred. No. 3.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TNYQRKRNIEDALNQLFRN 20
   |||||
Db 275 TNYQRKRNIEDALNQLFRN 294

```

Search completed: October 27, 2005, 18:44:24
 Job time : 70.2093 secs

; NUMBER OF SEQ ID NOS: 455

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 390
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-857-390

Query Match 100.0%; Score 105; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYORKNKRNIEDALNQLFRN 20
DB 275 TNYORKNKRNIEDALNQLFRN 294

RESULT 3
US-09-404-879A-389
; Sequence 389, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon B.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-389

Query Match 100.0%; Score 105; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYORKNKRNIEDALNQLFRN 20
DB 670 TNYORKNKRNIEDALNQLFRN 689

RESULT 4
US-09-667-857-389
; Sequence 389, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon B.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-857-389

Query Match 100.0%; Score 105; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYORKNKRNIEDALNQLFRN 20
DB 670 TNYORKNKRNIEDALNQLFRN 689

RESULT 5
US-09-404-879A-312
; Sequence 312, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon B.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-312

Query Match 100.0%; Score 105; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYORKNKRNIEDALNQLFRN 20
DB 751 TNYORKNKRNIEDALNQLFRN 770

RESULT 6
US-09-338-933-312
; Sequence 312, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-933-312

Query Match 100.0%; Score 105; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNYORKNKRNIEDALNQLFRN 20
DB 751 TNYORKNKRNIEDALNQLFRN 770

RESULT 7
US-09-667-857-312
; Sequence 312, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:

```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-667-857-312

```

```

Query Match          100.0%; Score 105; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TNYQRKNRIEDALNQLFRN 20
Db      751 TNYQRKNRIEDALNQLFRN 770

```

```

RESULT 8
US-09-404-879A-388
; Sequence 388, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 388
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-388

```

```

Query Match          65.7%; Score 69; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TNYQRKNRIEDA 13
Db      751 TNYQRKNRIEDA 763

```

```

RESULT 9
US-09-667-857-388
; Sequence 388, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.

```

```

; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 388
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-857-388

```

```

Query Match          65.7%; Score 69; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TNYQRKNRIEDA 13
Db      751 TNYQRKNRIEDA 763

```

```

RESULT 10
US-09-270-767-42105
; Sequence 42105, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 42105
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42105

```

```

Query Match          43.8%; Score 46; DB 4; Length 516;
Best Local Similarity 40.0%; Pred. No. 30;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 TNYQRKNRIEDALNQLFRN 20
Db      132 SNYRRIKQIMDILKMKFS 151

```

```

RESULT 11
US-09-248-796A-17941
; Sequence 17941, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17941
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Candida albicans

```

US-09-248-796A-17941

Query Match 43.3%; Score 45.5; DB 4; Length 336;

Best Local Similarity 64.7%; Pred. No. 22;

Matches 11; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 2 YORNRKNIEDALNQLF 18

Db 62 NYNNKRNRI---LNQWF 75

RESULT 12

US-09-115-150-5

Sequence 5, Application US/09115150

Patent No. 6503747

GENERAL INFORMATION:

APPLICANT: Kathariou, Sophia

APPLICANT: Lei, Xiang-Hei

TITLE OF INVENTION: Serotype-Specific Probes for *Listeria Monocytogenes*

FILE REFERENCE: A65378/RPT/DSS

CURRENT APPLICATION NUMBER: US/09/115,150

CURRENT FILING DATE: 1998-07-14

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 316

TYPE: PRT

ORGANISM: Bacterial

US-09-115-150-5

Query Match 42.9%; Score 45; DB 4; Length 316;

Best Local Similarity 50.0%; Pred. No. 25;

Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 TNYORNRKNIEDALNQLF 20

Db 121 TYTEBNKRNPHDFGNSLVN 140

RESULT 13

US-09-949-016-6984

Sequence 6984, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6984

LENGTH: 596

TYPE: PRT

ORGANISM: Human

US-09-949-016-6984

Query Match 42.4%; Score 44.5; DB 4; Length 596;

Best Local Similarity 64.7%; Pred. No. 61;

Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 YORNRKNIEDALNQLF 19

Db 554 YORNR-LEKALQOLAR 569

RESULT 14

US-09-949-016-7843

Sequence 7843, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7843

LENGTH: 632

TYPE: PRT

ORGANISM: Human

US-09-949-016-7843

Query Match 42.4%; Score 44.5; DB 4; Length 632;

Best Local Similarity 64.7%; Pred. No. 65;

Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 YORNRKNIEDALNQLF 19

Db 590 YORNR-LEKALQOLAR 605

RESULT 15

US-09-489-039A-7928

Sequence 7928, Application US/09489039A

Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *KLEBSIELLA*

TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7928

LENGTH: 334

TYPE: PRT

ORGANISM: *Klebsiella pneumoniae*

US-09-489-039A-7928

Query Match 41.9%; Score 44; DB 4; Length 334;

Best Local Similarity 53.3%; Pred. No. 38;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 YORNRKNIEDALNQL 17

Db 116 YRKNMSINDELSQL 130

Search completed: October 27, 2005, 17:34:32

Job time : 18.3953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:30:08 ; Search time 64 Seconds
(without alignments)
130.638 Million cell updates/sec

Title: US-10-612-090-15
Perfect score: 105
Sequence: 1 TTYORKKNIEDALNQLFRN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues
Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: Published Applications AA.*
 - 2: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
 - 3: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
 - 5: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
 - 7: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
 - 9: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubppa/US09C_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
 - 19: /cgn2_6/ptodata/2/pubppa/US11_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	20	15	US-10-612-090-15
2	105	100.0	108	15	US-10-612-090-2
3	105	100.0	108	15	US-10-612-090-13
4	105	100.0	178	9	US-09-884-441-489
5	105	100.0	178	10	US-09-907-969-489
6	105	100.0	178	14	US-10-198-053-489
7	105	100.0	178	17	US-10-860-790-489
8	105	100.0	233	9	US-09-884-441-488
9	105	100.0	233	10	US-09-907-969-488
10	105	100.0	233	14	US-10-198-053-488
11	105	100.0	233	17	US-10-860-790-488

12	105	100.0	284	10	US-09-965-738-300	Sequence 300, App
13	105	100.0	318	10	US-09-907-969-594	Sequence 594, App
14	105	100.0	318	14	US-10-198-053-554	Sequence 554, App
15	105	100.0	318	17	US-10-860-790-594	Sequence 594, App
16	105	100.0	367	15	US-10-333-900-27	Sequence 27, App
17	105	100.0	396	18	US-10-858-412-225	Sequence 225, App
18	105	100.0	438	9	US-09-884-441-483	Sequence 483, App
19	105	100.0	438	9	US-09-884-441-483	Sequence 483, App
20	105	100.0	438	10	US-09-907-969-390	Sequence 390, App
21	105	100.0	438	10	US-09-907-969-483	Sequence 483, App
22	105	100.0	438	10	US-09-827-271-350	Sequence 350, App
23	105	100.0	438	14	US-10-198-053-390	Sequence 390, App
24	105	100.0	438	14	US-10-333-900-19	Sequence 19, App
25	105	100.0	438	17	US-10-860-790-390	Sequence 390, App
26	105	100.0	438	17	US-10-860-790-390	Sequence 483, App
27	105	100.0	438	17	US-10-860-790-483	Sequence 483, App
28	105	100.0	439	10	US-09-965-738-148	Sequence 148, App
29	105	100.0	456	18	US-10-858-412-226	Sequence 226, App
30	105	100.0	526	15	US-10-333-900-30	Sequence 30, App
31	105	100.0	545	14	US-10-243-243A-4	Sequence 4, App
32	105	100.0	583	14	US-10-142-515-4	Sequence 4, App
33	105	100.0	748	17	US-10-687-035-1	Sequence 1, App
34	105	100.0	809	17	US-10-687-035-2	Sequence 1, App
35	105	100.0	833	9	US-09-884-441-389	Sequence 389, App
36	105	100.0	833	10	US-09-907-969-389	Sequence 389, App
37	105	100.0	833	10	US-09-827-271-389	Sequence 389, App
38	105	100.0	833	14	US-10-198-053-389	Sequence 389, App
39	105	100.0	833	17	US-10-860-790-389	Sequence 389, App
40	105	100.0	914	9	US-09-778-320-206	Sequence 206, App
41	105	100.0	914	9	US-09-910-689-206	Sequence 206, App
42	105	100.0	914	9	US-09-884-441-312	Sequence 312, App
43	105	100.0	914	9	US-09-884-441-478	Sequence 478, App
44	105	100.0	914	10	US-09-907-969-312	Sequence 312, App
45	105	100.0	914	10	US-09-907-969-478	Sequence 478, App

ALIGNMENTS

RESULT 1
US-10-612-090-15
; Sequence 15, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 15
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-15

Query Match 100.0%; Score 105; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTYORKKNIEDALNQLFRN 20
|||||
DB 1 TTYORKKNIEDALNQLFRN 20

RESULT 2
US-10-612-090-2
; Sequence 2, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:

US-10-198-053-489

Query Match 100.0%; Score 105; DB 14; Length 178;

Best Local Similarity 100.0%; Pred. No. 1.2e-07; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0;

QY 1 TNYQRKNRIEDALNQLFRN 20

DB 70 TNYQRKNRIEDALNQLFRN 89

RESULT 7

US-10-860-790-489

Sequence 489, Application US/10860790

Publication No. US20050031634A1

GENERAL INFORMATION:

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Retter, Marc W.

APPLICANT: Fanger, Gary R.

APPLICANT: Hill, Paul

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.462C11

CURRENT APPLICATION NUMBER: US/10/860,790

CURRENT FILING DATE: 2004-06-02

NUMBER OF SEQ ID NOS: 624

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 489

LENGTH: 178

TYPE: PRF

ORGANISM: Homo sapiens

US-10-860-790-489

Query Match 100.0%; Score 105; DB 17; Length 178;

Best Local Similarity 100.0%; Pred. No. 1.2e-07; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0;

QY 1 TNYQRKNRIEDALNQLFRN 20

DB 70 TNYQRKNRIEDALNQLFRN 89

RESULT 8

US-09-884-441-488

Sequence 488, Application US/09884441

Patent No. US20020119158A1

GENERAL INFORMATION:

APPLICANT: Algate, Paul A.

APPLICANT: Carter, Darrick

FILE REFERENCE: 210121.462C7

CURRENT APPLICATION NUMBER: US/09/884,441

CURRENT FILING DATE: 2001-06-18

NUMBER OF SEQ ID NOS: 489

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 488

LENGTH: 233

TYPE: PRF

ORGANISM: Homo sapiens

US-09-884-441-488

Query Match 100.0%; Score 105; DB 9; Length 233;

Best Local Similarity 100.0%; Pred. No. 1.6e-07; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0;

QY 1 TNYQRKNRIEDALNQLFRN 20

DB 70 TNYQRKNRIEDALNQLFRN 89

RESULT 9

US-09-907-969-488

Sequence 488, Application US/09907969

Publication No. US20030091580A1

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer L.

APPLICANT: King, Gordon E.

APPLICANT: Algate, Paul A.

APPLICANT: Fling, Steven P.

APPLICANT: Retter, Marc W.

APPLICANT: Fanger, Gary Richard

APPLICANT: Reed, Steven G.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Hill, Paul

APPLICANT: Albane, Earl

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.462C8

CURRENT APPLICATION NUMBER: US/09/907,969

CURRENT FILING DATE: 2001-07-17

NUMBER OF SEQ ID NOS: 596

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 488

LENGTH: 233

TYPE: PRF

ORGANISM: Homo sapiens

US-09-907-969-488

Query Match 100.0%; Score 105; DB 10; Length 233;

Best Local Similarity 100.0%; Pred. No. 1.6e-07; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0;

QY 1 TNYQRKNRIEDALNQLFRN 20

DB 70 TNYQRKNRIEDALNQLFRN 89

RESULT 10

US-10-198-053-488

Sequence 488, Application US/10198053

Publication No. US20030124140A1

GENERAL INFORMATION:

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Retter, Marc W.

APPLICANT: Fanger, Gary R.

APPLICANT: Hill, Paul

TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

FILE REFERENCE: 210121.462C9

CURRENT APPLICATION NUMBER: US/10/198,053

CURRENT FILING DATE: 2002-07-17

NUMBER OF SEQ ID NOS: 624

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 488

LENGTH: 233

TYPE: PRF

ORGANISM: Homo sapiens

US-10-198-053-488

Query Match 100.0%; Score 105; DB 14; Length 233;

Best Local Similarity 100.0%; Pred. No. 1.6e-07; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0;

QY 1 TNYQRKNRIEDALNQLFRN 20

DB 70 TNYQRKNRIEDALNQLFRN 89

RESULT 11

US-10-860-790-488

Sequence 488, Application US/10860790

Publication No. US20050031634A1

GENERAL INFORMATION:

APPLICANT: Bangur, Chaitanya S.

```

; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-790-488

Query Match          100.0%; Score 105; DB 17; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNYQRKNKRNIEDALNQLFRN 20
Db 70 TNYQRKNKRNIEDALNQLFRN 89

RESULT 12
US-09-965-738-300
; Sequence 300, Application US/09965738
; Publication No. US20030143667A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy
; TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic a
; FILE REFERENCE: 40715-258841
; CURRENT APPLICATION NUMBER: US/09/965,738
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/284,175
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 300
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-738-300

Query Match          100.0%; Score 105; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNYQRKNKRNIEDALNQLFRN 20
Db 121 TNYQRKNKRNIEDALNQLFRN 140

RESULT 13
US-09-907-969-594
; Sequence 594, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Pling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darwick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; AND DIAGNOSIS OF OVARIAN CANCER

```

```

; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 594
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-09-907-969-594

Query Match          100.0%; Score 105; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNYQRKNKRNIEDALNQLFRN 20
Db 155 TNYQRKNKRNIEDALNQLFRN 174

RESULT 14
US-10-198-053-594
; Sequence 594, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 594
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-10-198-053-594

Query Match          100.0%; Score 105; DB 14; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNYQRKNKRNIEDALNQLFRN 20
Db 155 TNYQRKNKRNIEDALNQLFRN 174

RESULT 15
US-10-860-790-594
; Sequence 594, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02

```



```

; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-10-860-790-594

```

```

Query Match          100.0%; Score 105; DB 17; Length 318;
Best Local Similarity 100.0%; Pred. No. 2,2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TNYQNRKNIEDALNQLFRN 20
DB 155 TNYQNRKNIEDALNQLFRN 174

```

Search completed: October 27, 2005, 18:31:46
Job time : 65 secs

(648) 742-0211

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 18:13:49 ; Search time 12.3721 Seconds
(without alignments)
155.538 Million cell updates/sec

Title: US-10-612-090-15

Perfect score: 105

Sequence: 1 TNYQRNKNIEDALNQLFRN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	48.6	841	2	T36384
2	48	45.7	666	2	T17396
3	47	44.8	1955	1	AGCH
4	46	43.8	535	2	S74703
5	46	43.8	1328	2	T43060
6	45	42.9	728	2	D84647
7	44	41.9	56	2	G81826
8	44	41.9	420	2	B72386
9	44	41.9	680	2	G82526
10	44	41.9	705	2	F70352
11	44	41.9	809	2	E90016
12	44	41.9	971	2	D70128
13	44	41.9	1879	2	T19481
14	44	41.4	475	2	H64218
15	43.5	41.4	2206	2	G71611
16	43	41.0	96	2	H70340
17	43	41.0	181	2	I40146
18	43	41.0	257	2	B70246
19	43	41.0	288	2	T35695
20	43	41.0	445	2	C71636
21	43	41.0	476	2	A28439
22	42	41.0	1240	2	T30834
23	42.5	40.5	94	2	H69887
24	42.5	40.5	469	2	T34173
25	42	40.0	188	2	A70256
26	42	40.0	201	2	E97049
27	42	40.0	278	2	D87517
28	42	40.0	297	2	C70251
29	42	40.0	302	2	S50609

30	42	40.0	306	2	S61169	hypothetical prote
31	42	40.0	334	2	D86892	hypothetical prote
32	42	40.0	343	2	C55590	hypothetical prote
33	42	40.0	384	2	H89873	hypothetical prote
34	42	40.0	433	2	UC1230	DNA-binding prote
35	42	40.0	510	2	AG1539	oligo-1,6-glucosid
36	42	40.0	571	2	S58356	pept protein - Sta
37	42	40.0	605	2	S64410	probable membrane
38	42	40.0	680	2	C82905	hypothetical prote
39	42	40.0	906	2	B96901	uncharacterized co
40	42	40.0	1214	2	T21915	hypothetical prote
41	42	40.0	1279	2	E64709	type IIS restricti
42	42	40.0	2269	2	T28677	thoxy protein -
43	41	39.0	108	2	T27649	hypothetical prote
44	41	39.0	225	2	T44720	hypothetical prote
45	41	39.0	232	2	H87113	conserved hypotet

ALIGNMENTS

RESULT 1
T36384
probable ATP-binding proteinase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36384
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A:Reference number: Z21573
A:Accession: T36384
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-841
A:Cross-references: UNIPROT:Q9SGT8; EMBL:AL049628; PIDN:CAB40873.1; GSPDB:GND0070; SCODE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODEB:SC94.24C
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: ATP; molecular chaperone; nucleotide binding; P-loop
F:212-219/Region: nucleotide-binding motif A (P-loop)
F:279-284/Region: nucleotide-binding motif B
F:549-556/Region: nucleotide-binding motif A (P-loop)
F:617-622/Region: nucleotide-binding motif B
F:218/Binding site: ATP (lys) #status predicted
F:555/Binding site: ATP (lys) #status predicted

Query Match 48.6%; Score 51; DB 2; Length 841;
Best Local Similarity 47.4%; Pred. No. 13;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TNYQRNKNIEDALNQLFR 19
Db 685 SNYERKNKVDDELKQHR 703

RESULT 2
T17396
vlp protein - Dichelobacter nodosus
C:Species: Dichelobacter nodosus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T17396
R:Billington, S.J.; Huggins, A.S.; Johansen, P.A.; Crellin, P.K.; Cheung, J.K.; Katz, M.
Infect. Immun. 67, 1277-1286, 1999
A:Title: Complete nucleotide sequence of the 27-kilobase virulence related locus (vrl) of
A:Reference number: Z18734; MUID:99150261; PMID:10024571
A:Accession: T17396
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-666 <BL>
A:Cross-references: UNIPROT:Q32495; EMBL:U20246; NID:g3493323; PID:g2317809; PIDN:AAC333
A:Experimental source: strain A198
C:Superfamily: Dichelobacter nodosus vlp protein

Query Match 45.7% Score 48; DB 2; Length 666;
 Best Local Similarity 55.6%; Pred. No. 29;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TNYORKNKINIDALNQLF 18
 DB 345 TBYOTRREIRGIDALNTIF 362

RESULT 3
 AGCH
 agrin precursor - chicken
 CSpecies: Gallus gallus (chicken)
 CDate: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
 CAccession: JH0591; A38857; B38857; I50692
 R/Tsrm, K.W.K.; Ruegg, M.A.; Escher, G.; Kroege, S.; McMahon, U.J.
 Neuron 8, 677-689, 1992
 A>Title: cDNA that encodes active agrin.
 AReference number: JH0591; MUID:92232297; PMID:1314620
 AAccession: JH0591
 A.Molecule type: mRNA
 A.Residues: 1-1955 <TS1>
 A.Cross-references: UNIPROT:P31696; GB:M94271; NID:g211120; PIDN:AAA48565.1; PID:g211121
 A.Experimantal source: brain
 R/Ruegg, M.A.; Tsrm, K.W.K.; Horton, S.E.; Kroege, S.; Escher, G.; Gensch, E.M.; McMahon
 Neuron 8, 691-699, 1992
 A>Title: The agrin gene codes for a family of basal lamina proteins that differ in funct
 AReference number: A38857; MUID:92232298; PMID:1314621
 A.Contents: alternative splicing
 AAccession: A38857
 A.Molecule type: mRNA
 A.Residues: 1132-1783; 1795-1955 <R2>
 A.Cross-references: GB:M97371
 AAccession: B38857
 A.Molecule type: mRNA
 A.Residues: 1221-1647; 1653-1783; 1794-1955 <R3>
 A.Cross-references: GB:M97372
 A.Note: translation of the nucleotide sequence is not complete
 R/Thomas, W.S.; O'Dowd, D.K.; Smith, M.A.
 Dev. Biol. 158, 523-535, 1993
 A>Title: Developmental expression and alternative splicing of chick agrin RNA.
 AReference number: I50692; MUID:93345745; PMID:8393816
 AAccession: I50692
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 'SHLSNEIPA', 1784-1795 <THO>
 A.Cross-references: EMBL:U07271; NID:g459665; PIDN:AAA16788.1; PID:g459666
 CComment: This protein mediates the motor neuron-induced aggregation of acetylcholine r
 CSuperfamily: agrin; EGF homology; Kazal proteinase inhibitor homology; laminin G repe
 CKeywords: alternative splicing; duplication; glycoprotein; neuromuscular junction
 F.1-38/Domain: signal sequence #status predicted <SIG>
 F.39-1955/Product: agrin #status predicted <MAT>
 F.39-1783; 1795-1955/Product: agrin-related protein 1 #status predicted <AG1>
 F.39-1647; 1652-1783; 1794-1955/Product: agrin-related protein 2 #status predicted <AG2>
 F.177-126/Domain: Kazal proteinase inhibitor homology <KP11>
 F.152-201/Domain: Kazal proteinase inhibitor homology <KP12>
 F.125-273/Domain: Kazal proteinase inhibitor homology <KP13>
 F.1295-344/Domain: Kazal proteinase inhibitor homology <KP14>
 F.3370-418/Domain: Kazal proteinase inhibitor homology <KP15>
 F.435-483/Domain: Kazal proteinase inhibitor homology <KP16>
 F.500-548/Domain: Kazal proteinase inhibitor homology <KP17>
 F.584-633/Domain: Kazal proteinase inhibitor homology <KP18>
 F.675-726/Domain: laminin-type EGF-like homology <LE1>
 F.729-773/Domain: laminin-type EGF-like homology <LE2>
 F.801-851/Domain: Kazal proteinase inhibitor homology <KP19>
 F.856-995/Region: serine/threonine-rich
 F.1150-1219/Region: serine/threonine-rich
 F.1233-1264/Domain: EGF homology <EG1>
 F.1294-1448/Domain: laminin G repeat homology <LG1>
 F.1429-1431/Region: motor neuron attachment (L-R-E) motif
 F.1450-1482/Domain: EGF homology <EG2>

F.1489-1521/Domain: EGF homology <EG3>
 F.1560-1711/Domain: laminin G repeat homology <LG2>
 F.1718-1751/Domain: EGF homology <EG4>
 F.1803-1955/Domain: laminin G repeat homology <LG3>
 F.86-105; 94-126; 160-180; 169-201; 233-252; 241-273; 304-323; 312-344; 378-397; 386-418; 443-462;
 1482; 1489-1500; 1494-1510; 1512-1521/Duplicate bonds: #status predicted
 F.390; 659; 764; 814/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 44.8% Score 47; DB 1; Length 1955;
 Best Local Similarity 69.2%; Pred. No. 1.2e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 RNIEDALNQLFRN 20
 DB 1063 RSISBALDELFRN 1075

RESULT 4
 574703
 hypothetical protein slr1301 - Synechocystis sp. (strain PCC 6803)
 CSpecies: Synechocystis sp.
 A.Variety: PCC 6803
 CDate: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 CAccession: S74703
 R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
 DNA Res. 3, 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.
 AReference number: S74322; MUID:97061201; PMID:8905231
 AAccession: S74703
 A>Status: nucleic acid sequence not shown; translation not shown
 A.Molecule type: DNA
 A.Residues: 1-535 <KAN>
 A.Cross-references: UNIPROT:P72839; EMBL:D90901; GB:AB001339; NID:G1651897; PIDN:BAAL6854
 A.Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C.Genetics:
 A.Start codon: GTG

Query Match 43.8% Score 46; DB 2; Length 535;
 Best Local Similarity 43.8%; Pred. No. 45;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TNYORKNKINIDALNQ 16
 DB 481 TNYEONRREVEDLOKQ 496

RESULT 5
 743060
 agrin - electric ray (Discopyge ommata) (fragment)
 CSpecies: Discopyge ommata
 CDate: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 CAccession: T43060
 R/Smith, M.A.; Magill, S.C.; Rupp, F.; Yao, Y.M.M.; Schilling, J.W.; Snow, P.; McMahar
 submitted to the EMBL Data Library, September 1992
 AReference number: Z22308
 AAccession: T43060
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-1328 <SMI>
 A.Cross-references: UNIPROT:Q90404; EMBL:L01423; NID:g213102; PID:g213103; PIDN:AAA49224.
 CSuperfamily: agrin; EGF homology; Kazal proteinase inhibitor homology; laminin G repeat
 CKeywords: glycoprotein; neuromuscular junction

Query Match 43.8% Score 46; DB 2; Length 1328;
 Best Local Similarity 69.2%; Pred. No. 1.2e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 RNIEDALNQLFRN 20
 DB 448 RSIEBALDELFRN 460

RESULT 6
D84647
hypothetical protein AC2G25350 [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: D84647
R/Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84647
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-728 <STO>
A/Cross-references: GB:AE002093; NID:96598555; PIDN:AAD20696.2; GSPDB:GN00139
A/Genetic: A:Gene: AC2G25350
A/Map position: 2

Query Match 42.9%; Score 45; DB 2; Length 728;
Best Local Similarity 64.3%; Pred. No. 88;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TNYQRKNRDNALQ 14
| | | | | : | | | | |
Db 386 TTKQRKNRDNALQ 399

RESULT 7
G81826
hypothetical protein NMA1978 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup
C/Species: *Neisseria meningitidis*
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: G81826
R/Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A/Reference number: A81775; MUID:2022556; PMID:10761919
A/Accession: G81826
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-56 <PAR>
A/Cross-references: UNIPROT:Q9J743; GB:AL162757; GB:AL157959; NID:97380371; PIDN:CAB8519
A/Experimental source: serogroup A, strain Z2491
C/Genetic: A:Gene: NMA1978

Query Match 41.9%; Score 44; DB 2; Length 56;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 ORRKNRDNALQ 18
: | | | | : | | | | : | | | |
Db 28 RRRKNRDNALQ 42

RESULT 8
B72386
hypothetical protein - *Thermotoga maritima* (strain MSB8)
C/Species: *Thermotoga maritima*
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: B72386
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A/Reference number: A72200; MUID:99287316; PMID:10360571
A/Accession: B72386

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-420 <ARN>
A/Cross-references: UNIPROT:Q9WY19; GB:AE001716; GB:AE000512; NID:94980853; PIDN:AAD35441
A/Experimental source: strain MSB8
C/Genetic: A:Gene: TM0354

Query Match 41.9%; Score 44; DB 2; Length 420;
Best Local Similarity 57.1%; Pred. No. 70;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQRKNRDNALQ 16
| | | | | : | | | | |
Db 318 YQEVVKNRDNALQ 331

RESULT 9
G82526
ATP-dependent DNA helicase XP2680 [imported] - *Xylella fastidiosa* (strain 9a5c)
C/Species: *Xylella fastidiosa*
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: G82526
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: G82526
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-680 <SIM>
A/Cross-references: UNIPROT:Q9PA40; GB:AE004074; GB:AE003849; NID:91079118; PIDN:AFR8547.

A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Bixones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H.
submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E.
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanelli, R.V.; Sawasaki
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsunako, M.H.; Vailada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze
A/Reference number: A59328
A/Contents: annotation
C/Genetic: A:Gene: XP2680
C/Superfamily: helicase II

Query Match 41.9%; Score 44; DB 2; Length 680;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NYQRKNRDNALQ 12
: | | | | : | | | | : | | | |
Db 518 SYQRKNRDNALQ 528

RESULT 10
F70352
conserved hypothetical protein aq_585 - *Aquifex aeolicus*
C/Species: *Aquifex aeolicus*
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: F70352
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: F70352

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-705 <AOF>
A:Cross-references: UNIPROT:O66850; GB:AEO00697; NID:g2983212; PIDN:AAC06820.1; PID:g2983212
A:Experimental source: strain VFS
A:Genetics:
A:Gene: ag_585

Query March 41.9%; Score 44; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 YORNRKNIEDALNQLFRN 20
| : | | | : | | | : |
Db 303 YFKKKRFDLEIALSRKEN 320

RESULT 11
E90016
hyaluronate lyase precursor [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E90016
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogudma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: E90016
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-809 <KUR>
A:Cross-references: UNIPROT:Q99567; GB:BA000018; PID:g13702002; PIDN:BAB43294.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: hyaA

Query March 41.9%; Score 44; DB 2; Length 809;
Best Local Similarity 42.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 NYORKRNIEDALNQLFRN 20
| | : : | | | : | |
Db 136 NTDENKKVKDLEIWRKN 154

RESULT 12
D70128
conserved hypothetical protein BB0228 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: D70128
R:Fischer, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitehouse, D.; Peterson, J.; Karlavag, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugrson, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: AF0100; MUID:98065943; PMID:9403685
A:Accession: D70128
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-971 <KLE>
A:Cross-references: UNIPROT:O51246; GB:AEO01133; GB:AEO00783; NID:g2688120; PIDN:AAC6662
A:Experimental source: strain BJ1

Query Match 41.9%; Score 44; DB 2; Length 971;
Best Local Similarity 47.4%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 NYORKRNIEDALNQLFRN 20
| : | | | : | | | : |

```

DB          896  SYRRKMLNISDSLRODIRN 914

RESULT 13
T19481
Hypochemical protein C26C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19481; T25434
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19128
A:Accession: T19481
A:Status: preliminary; translated from GB/EMBL/DBDJI
A:Molecule type: DNA
A:Residues: 1-1879 <W1>
A:Cross-references: UNIPROT:Q18210; EMBL:Z72503; PIDN:CAA96600.1; GSPDB:GN00019; CESP:C26C6
A:Experimental source: clone C26C6
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z20033
A:Accession: T25434
A:Status: preliminary; translated from GB/EMBL/DBDJI
A:Molecule type: DNA
A:Residues: 1-1879 <W12>
A:Cross-references: EMBL:Z72517; PIDN:CAA96697.1; GSPDB:GN00019; CESP:C26C6.1
A:Experimental source: clone T28F4
C:Genetics:
A:Gene: CESP:C26C6.1
A:Map position: 1
A:Introns: 42/3; 295/2; 375/2; 489/3; 665/2; 980/3; 1233/2; 1292/1; 1325/3; 1349/3; 1486/3
F:368-422/Domain: bromodomain homology <BRO1>
F:560-615/Domain: bromodomain homology <BRO2>
F:737-792/Domain: bromodomain homology <BRO3>

Query Match          41.9%; Score 44; DB 2; Length 1879;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

CY          2  NYQNRKNRIEDALNOL 17
| : : : : : : : : : :
DB          692  NIQHQRSMEDALIMOL 707

RESULT 14
H64218
preprotein translocase secY - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
C:Accession: H64218
R:Risser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M. ; Fuhrman, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.R.; C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MID:96026346; PMID:7569993
A:Accession: H64218
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-475 <T1GR>
A:Cross-references: UNIPROT:P47416; GB:U39695; GB:L43967; NID:g1045833; PID:g1045854; T19481
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: preprotein translocase secY

Query Match          41.4%; Score 43.5; DB 2; Length 475;
Best Local Similarity 38.5%; Pred. No. 95;
Matches 10; Conservative 5; Mismatches 2; Indels 9; Gaps 1;

CY          1  TNYQNRKNRI-----EDALNOL 17
| : : : : : : : : : :
DB          439  TNYQQLRRNLAIIEVQOTACDLSLEOL 464

```

RESULT 15

G71611
 hypothetical protein PFB0560w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C:Accession: G71611
 R:Gardner, M.J.; Tetteijn, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 .; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
 Science 282, 1126-1132, 1998
 A>Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; WUID:99021743; PMID:9804551
 A:Accession: G71611
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2206 <GAR>
 A:Cross-references: GB:AE001403; GB:AE001362; NID:G3845216; PIDN:AACT1901.1; PID:G384521
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0560w

Query Match 41.4%; Score 43.5; DB 2; Length 2206;
 Best Local Similarity 36.8%; Pred. No. 4.7e+02;
 Matches 7; Conservative 6; Mismatches 5; Indels 1; Gaps 1;
 QY 2 NYQRNKRNI-DALNQLFR 19
 || | |::| |:::
 Db 1972 NYNNNNNVDDVLYNKIKYK 1990

Search completed: October 27, 2005, 18:57:25
 Job time : 14.3721 secs

This Page Blank (intentional)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:34:49 ; Search time 58.1395 Seconds
(without alignments)
176.155 Million cell updates/sec

Title: US-10-612-090-15
Perfect score: 105
Sequence: 1 TNYQRNKNIEDALNQLFRN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	100.0	1148	2 Q9H7S7	Q9H7S7 homo sapien
2	105	100.0	6995	2 Q96RK2	Q96RK2 homo sapien
3	105	100.0	22152	2 Q8WY17	Q8WY17 homo sapien
4	79	75.2	258	2 Q9DHI1	Q9DHI1 mus musculu
5	54	51.4	482	2 Q7QDT5	Q7QDT5 anopheles g
6	52	49.5	390	2 Q7RCC0	Q7RCC0 plasmodium
7	51	48.6	841	2 Q9S6T8	Q9S6T8 streptomyce
8	48	45.7	109	2 Q9XZ81	Q9XZ81 plasmodium
9	48	45.7	132	2 Q6FCV7	Q6FCV7 acinetobact
10	48	45.7	319	2 Q8ZC8	Q8ZC8 salmonella
11	48	45.7	360	2 Q7SB96	Q7SB96 neurospora
12	48	45.7	666	2 Q32495	Q32495 bacteroides
13	48	45.7	708	2 Q9CQ16	Q9CQ16 manduca sex
14	48	45.7	797	2 Q8KMH3	Q8KMH3 microomosp
15	48	45.7	841	2 Q8ZEB8	Q8ZEB8 streptomyce
16	48	45.7	956	2 Q6LQD0	Q6LQD0 photobacter
17	48	45.7	2879	2 Q6LFO9	Q6LFO9 plasmodium
18	47	44.8	349	2 Q7NFA9	Q7NFA9 gloeobacter
19	47	44.8	447	2 Q6W3Q4	Q6W3Q4 alvineella p
20	47	44.8	575	2 Q812A1	Q812A1 plasmodium
21	47	44.8	1048	2 Q8KJ11	Q8KJ11 rhicobium l
22	47	44.8	1955	1 AGRN_CHICK	P31696 gallus gall
23	46.5	44.3	2060	2 Q7Q3J9	Q7Q3J9 anopheles g
24	46	43.8	109	2 Q9NHJ1	Q9NHJ1 plasmodium
25	46	43.8	171	1 PACA_RANRI	Q9N169 r glucagon-
26	46	43.8	376	2 Q8PMN7	Q8PMN7 methanosarc
27	46	43.8	535	2 P72839	P72839 synecocyst
28	46	43.8	921	2 Q7RP27	Q7RP27 plasmodium
29	46	43.8	970	2 Q7RG16	Q7RG16 plasmodium
30	46	43.8	1049	2 Q83611	Q83611 ectromelia
31	46	43.8	1049	2 Q83611	Q83611 ectromelia

32	46	43.8	1113	2 Q8JL89	Q8JL89 ectromelia
33	46	43.8	1328	1 AGRN DISOM	Q90404 discopyge o
34	46	43.8	2760	2 Q815Y2	Q815Y2 plasmodium
35	45.5	43.3	792	2 Q6R230	Q6R230 debaryomyce
36	45.5	43.3	992	2 Q99877	Q99877 trypanosoma
37	45	42.9	109	2 Q8T4V4	Q8T4V4 plasmodium
38	45	42.9	110	2 Q931B6	Q931B6 staphylococ
39	45	42.9	247	2 Q7PVY8	Q7PVY8 anopheles g
40	45	42.9	248	2 Q8IKY4	Q8IKY4 plasmodium
41	45	42.9	314	2 Q8WUT3	Q8WUT3 tetrahymena
42	45	42.9	315	2 Q9ZTCS	Q9ZTCS listeria mo
43	45	42.9	316	2 Q71W10	Q71W10 listeria mo
44	45	42.9	386	2 Q81DW4	Q81DW4 plasmodium
45	45	42.9	447	2 Q6W3M3	Q6W3M3 alvineella p

ALIGNMENTS

RESULT 1
Q9H7S7 PRELIMINARY; PRT; 1148 AA.
ID Q9H7S7
AC Q9H7S7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; DOI=10.1038/ng1285;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Iehi S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahori K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiroka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottura T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Moriya S., Motiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamagaki M., Watanabe K., Kumagai A., Itakura S., Fukunumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Taniguchi A., Fujiiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunaga K., Nakajima Y., Mizuno T., Morinaga M., Saeaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK024365; BAB14899.1; -.
DR HSSP; Q9DHI1; IIVZ.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 7.
DR PROSITE; PS0024; SEA; 3.
SQ SEQUENCE 1148 AA; 127957 MW; 3861B0D5EFPBABC CRC64;

```

Query Match          100.0%; Score 105; DB 2; Length 1148;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 TNYQNRKNRNIEDALNQLFRN 20
    |||||
    985 TNYQNRKNRNIEDALNQLFRN 1004

RESULT 2
ID Q96RK2 PRELIMINARY; PRT; 6995 AA.
AC Q96RK2
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DR 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DN Mucin 16 (Fragment).
GN Name=MUC16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200;
RY Yln B.W., Lloyd K.O.;
RT "Molecular cloning of the cal25 ovarian cancer antigen. Identification
RT as a new mucin, mucle."
RL J. Biol. Chem. 276:27371-27375(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Lloyd K.O., Yln B.W.T.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF361486; AAK74120.3; -.
DR HSSP; Q9DIH1; 11VZ.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 20.
DR SMART; SMO0200; SEA; 10.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00024; SEA; 6.
RT NON TER
SO SEQUENCE 6995 AA; 744958 MW; 80C797DBDF33A2B CRC64;

Query Match          100.0%; Score 105; DB 2; Length 6995;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 TNYQNRKNRNIEDALNQLFRN 20
    |||||
    6832 TNYQNRKNRNIEDALNQLFRN 6851

RESULT 3
ID Q8WX17 PRELIMINARY; PRT; 22152 AA.
AC Q8WX17;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DR 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DN Ovarian cancer related tumor marker CA125.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
RY O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RY York L.;
RT "The CA 125 gene: an extracellular superstructure dominated by repeat
RT sequences."
RL Tumour Biol. 22:348-366(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RY O'Brien T.J., Underwood L.J., Beard J.B.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414442; AAL65133.2; -.
DR Genew; HGNC:15582; MUC16.
DR Pfam; PF01390; SEA; 51.
DR SMART; SMO0200; SEA; 23.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_2.
DR PROSITE; PS00024; SEA; 11.
SQ SEQUENCE 22152 AA; 235268 MW; B3E7BDF1997A440 CRC64;

Query Match          100.0%; Score 105; DB 2; Length 22152;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 TNYQNRKNRNIEDALNQLFRN 20
    |||||
    21989 TNYQNRKNRNIEDALNQLFRN 22008

RESULT 4
ID Q9DIH1 PRELIMINARY; PRT; 258 AA.
AC Q9DIH1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DR 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:111000814 product:hypothetical SEA domain containing
DE protein, full insert sequence.
DN Name=111000814R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carinici P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carinici P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carinici P.,
RA Kono H., Akiyama J., Nishi K., Kitsumi T., Tashiro H., Itoh M.,

```

DR	InterPro; IPRO01320; Ion_glu_receptor.
DR	InterPro; IPR001622; K+channel_pore.
DR	InterPro; IPR001508; NMDA_receptor.
DR	InterPro; IPRO01311; SBP/Glu_receptor.
DR	Pfam; PF00060; Lig_chan; 1.
DR	PRINTS; PRO0177; NMDARECEPTOR.
FT	NON_TER 1
FT	NON_TER 482
SQ	SEQUENCE 482 AA; 54936 MW; 2BB7187196C8DFE7 CRC64;
Query Match	51.4%; Score 54; DB 2; Length 482;
Best Local Similarity	52.6%; Pred. No. 11;
Matches 10; Conservative	4; Mismatches 5; Indels 0; Gaps 0;
OY	2 NYORKNKIEDALNQFRN 20 :: : ::
Dd	272 DYFNMGKADVLVQMWRN 290
RESULT 6	
O7RCO	PRELIMINARY; PRT; 390 AA.
ID	O7RCO
AC	O7RCO;
DT	01-MAR-2004 (TReMBLrel_13, Created)
DT	01-MAR-2004 (TReMBLrel_26, Last sequence update)
DT	01-MAR-2004 (TReMBLrel_26, Last annotation update)
DE	Hypothetical protein.
CN	Name=PY05864;
OS	Plasmodium yoelii yoelii.
CC	Eukaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=73239;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=17XNL;
EX	PubMed=12368865; DOI=10.1038/nature01099;
RA	Carlton J.M., Angiolini S.V., Sun B.B., Koczi T.W., Perlea M., Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L., Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L., Shallom S.J., van Aken S.E., Riedmiller S.B., Feldlyum T.V., Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M., Florens L., Yates F.R. III, Raine J.D., Sindén R.E., Harris M.A., Rammungham D.A., Preiser P.R., Bergman L.W., Valdaya A.B., van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R., Ra Salzborg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J., Carucci D.J.;
RT	"Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";
RL	Nature 419:512-519(2002).
CC	-!- CAUTION: The sequence shown here is derived from an EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
CC	EMBL; AABL01001920; EAA17959.1; -.
DR	InterPro; IPR011591; Botulinum.
DR	ProDom; PD001963; Botulinum; 1.
KW	Hypothetical protein.
SQ	SEQUENCE 390 AA; 45937 MW; 0FA3B215E0A68BEA CRC64;
Query Match	49.5%; Score 52; DB 2; Length 390;
Best Local Similarity	60.0%; Pred. No. 18;
Matches 9; Conservative	3; Mismatches 3; Indels 0; Gaps 0;
OY	2 NYORKNKIEDALNQ 16 : ::
Dd	233 NYEKKNENIENILQQ 247
RESULT 7	
O9S6T8	PRELIMINARY; PRT; 841 AA.
ID	O9S6T8
AC	O9S6T8;
DT	01-MAY-2000 (TReMBLrel_13, Created)
DT	01-MAY-2000 (TReMBLrel_13, Last sequence update)
DT	01-OCT-2003 (TReMBLrel_25, Last annotation update)

```

DE Putative C1P-family ATP-binding protease.
GN ORFNames=SCB94.24c;
GN Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RT Nature 417:141-147(2002).
DR EMBL: AL939116; CAB40873.1; -.
DR PIR: T36384; T36384.
DR HSSP: P03815; IJBK.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004518; F:nuclease activity; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0006289; F:nucleotide-excision repair; IEA.
DR GO: GO:0019538; P:protein metabolism; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001270; Chaprln_C1PA/B.
DR InterPro: IPR004176; C1P_N.
DR InterPro: IPR001943; UVRB/C.
DR Pfam: PF00004; AAA; 1.
DR Pfam: PF02861; C1P_N; 2.
DR Pfam: PF02151; UVR_1.
DR PRINTS: PR00300; CLPROTEASEA.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS50151; UVR; 1.
DR ATP-binding; Complete proteome; Protease.
KW SEQUENCE 841 AA; 93013 MW; 101BE09AADA7FFD CRC64;
SQ
Query Match 48.6%; Score 51; DB 2; Length 841;
Best Local Similarity 47.4%; Pred. No. 57;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
OY 1 TNYQNRKNIEDALNQLFR 19
DB 685 SNYERMKRVQDELKQHR 703
RESULT 8
OX XZ81 PRELIMINARY; PRT; 109 AA.
AC O9XZ81;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DBL alpha protein (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=20078956; PubMed=10613695; DOI=10.1016/S0166-6851(99)00159-0;
RA Taylor H.M., Kyes S.A., Harris D., Kriek N., Newbold C.I.;

```

```

RT "A study of var gene transcription in vitro using universal var gene
RT primers.";
RT Mol. Biochem. Parasitol. 105:13-23(2000).
DR EMBL: AF133853; AAD33614.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 109 AA; 13027 MW; 7EFD9E0BD144E174 CRC64;
Query Match 45.7%; Score 48; DB 2; Length 109;
Best Local Similarity 44.4%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
OY 3 YQNRKNIEDALNQLFRN 20
DB 15 HQQRKRLKENLKQMFEN 32
RESULT 9
OX FCV7 PRELIMINARY; PRT; 132 AA.
AC O6FCV7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=ACIAD1226;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallenet D., Fonknechten N., Kreilmeyer A., Ozas S.,
RA Labrie L., Crueviller S., Robert C., Duprat S., Wincker P.,
RA Orston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
DR EMBL: CR543861; CAG68102.1; -.
KW Complete proteome.
SQ SEQUENCE 132 AA; 14775 MW; E8B481E44AB30F2A CRC64;
Query Match 45.7%; Score 48; DB 2; Length 132;
Best Local Similarity 56.2%; Pred. No. 23;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 2 NYQNRKNIEDALNQL 17
DB 80 DYQNRKNIEDALNQL 95
RESULT 10
OX ZOC8 PRELIMINARY; PRT; 319 AA.
AC O8ZOC8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Homologous to secreted protein soppd.
GN OrderedLocusNames=STM0972;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvarey E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;

```



```

DT 01-OCT-2002 (TRENBLrel. 22, last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, last annotation update)
DR CaIR4.
GN Name=caIR4;
OS Micromonospora echinospora (Micromonospora purpurea).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micromonosporaceae; Micromonosporaceae; Micromonospora.
CX NCBI_TaxID=1877;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 15839;
RX MEDLINE=22171414; PubMed=12183629; DOI=10.1126/science.1072105;
RA Ahlert J., Shepard E., Lomovskaya N., Zaopoulos E., Staffa A.,
RA Bachmann B.O., Huang K., Fongstein L., Czisny A., Whitlam R.E.,
RA Farnet C.M., Thorson J.S.;
RT "The calicheamicin gene cluster and its iterative type I enediylase
RT PKS.";
RL Science 287:1173-1176(2002).
DR EMBL; AF487482; AAM94782.1; -.
DR HSSP; P03815; IUBK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; F:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001270; Chaprinin_c1pA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; Uvr; 1.
DR PRINTS; PR00300; CLPPROTEASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS50151; UVR; 1.
DR ATP-binding.
SQ SEQUENCE 797 AA; 88267 MW; 9E107A6CA894C190 CRC64;

Query Match 45.7%; Score 48; DB 2; Length 797;
Best Local Similarity 47.4%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TNYQRNKRNIEDALNQLFR 19
DB 648 SNYDRMKQKVNDELKQHR 666

RESULT 15
Q82EB8 PRELIMINARY; PRT; 841 AA.
AC Q82EB8;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, last annotation update)
DE Putative ATP-dependent Clp protease.
GN Name=clpC2; Order=clpC2; Names=SAVA697;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
CX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shida T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shida T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AP005040; BAC72409.1; -.
DR HSSP; P03815; IUBK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001270; Chaprinin_c1pA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; PR00300; CLPPROTEASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS50151; UVR; 1.
DR ATP-binding; Complete proteome.
SQ SEQUENCE 841 AA; 93023 MW; 265866F0CA9BCCF7 CRC64;

Query Match 45.7%; Score 48; DB 2; Length 841;
Best Local Similarity 47.4%; Pred. No. 1.6e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 TNYQRNKRNIEDALNQLFR 19
DB 685 SNYERMKKVNDELKQHR 703

```

Search completed: October 27, 2005, 18:55:10
Job time : 61.1395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:31:33 ; Search time 72.6698 Seconds
(without alignments)
111.765 Million cell updates/sec

Title: US-10-612-090-16

Perfect score: 107
Sequence: 1 FRNSSIKSYFSDQVSTFRSV 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	107	100.0	21	8	ADJ66743	Adj66743 Human Muc
2	107	100.0	108	8	ADJ66729	Adj66729 Human Muc
3	107	100.0	108	8	ADJ66740	Adj66740 Human Muc
4	107	100.0	174	7	ADM30756	Adm30756 Human can
5	107	100.0	178	5	ABP30979	Abp30979 Predicted
6	107	100.0	178	7	ADA08642	Ada08642 Human O77
7	107	100.0	178	7	ADF08985	Adf08985 Secreted
8	107	100.0	233	5	ABP30978	Abp30978 Extracell
9	107	100.0	233	7	ADA08641	Ada08641 Human O77
10	107	100.0	233	7	ADF08984	Adf08984 Secreted
11	107	100.0	284	6	ABU54859	Abu54859 Human CA1
12	107	100.0	318	5	ABP31025	Abp31025 Amino aci
13	107	100.0	318	7	ADA08747	Ada08747 Human O77
14	107	100.0	318	7	ADF09090	Adf09090 Secreted
15	107	100.0	367	4	AAE12631	Aae12631 Human gen
16	107	100.0	438	3	AAE12555	Aae12555 Human ova
17	107	100.0	438	4	AAE12623	Aae12623 Human gen
18	107	100.0	438	5	ABP30973	Abp30973 Partial p
19	107	100.0	438	5	ABP30899	Abp30899 0772P c10
20	107	100.0	438	7	ADA08636	Ada08636 Human O77
21	107	100.0	438	7	ADA08543	Ada08543 Human ova
22	107	100.0	438	7	ADP08886	Adp08886 Secreted
23	107	100.0	438	7	ADP08979	Adp08979 Secreted
24	107	100.0	438	7	ADG46173	Adg46173 Human ova
25	107	100.0	439	6	ABU54709	Abu54709 Human CA1

26	107	100.0	526	4	AAE12634	Aae12634 Human gen
27	107	100.0	748	8	ADS94302	Ads94302 CA 125/O7
28	107	100.0	809	8	ADS94303	Ads94303 CA 125/O7
29	107	100.0	833	3	AAE12554	Aae12554 Human ova
30	107	100.0	833	5	ABP30898	Abp30898 0772P c10
31	107	100.0	833	7	ADA08542	Ada08542 Human ova
32	107	100.0	833	7	ADP08885	Adp08885 Secreted
33	107	100.0	833	7	ADG46172	Adg46172 Human ova
34	107	100.0	914	3	AAE12552	Aae12552 Human ova
35	107	100.0	914	4	AAE12553	Aae12553 Human ova
36	107	100.0	914	5	ABP30968	Abp30968 Hypothetcl
37	107	100.0	914	5	ABP30896	Abp30896 0772P p10
38	107	100.0	914	7	ADA08631	Ada08631 Human O77
39	107	100.0	914	7	ADA08465	Ada08465 Human ova
40	107	100.0	914	7	ADF08974	Adf08974 Secreted
41	107	100.0	914	7	ADF08808	Adf08808 Secreted
42	107	100.0	914	7	ADG46095	Adg46095 Human ova
43	107	100.0	914	8	ADN40451	Adn40451 Human bre
44	107	100.0	1148	4	AAE12583	Aae12583 Human pro
45	107	100.0	1148	4	ABE50283	ABE50283 HOST-1 ov

ALIGNMENTS

RESULT 1

ADJ66743
ID ADJ66743 standard; peptide; 21 AA.

AC ADJ66743;
XX
DT 06-MAY-2004 (first entry)
XX

DE Human Muc1/Muc16 protein-related peptide SegID16.

XX monoclonal antibody; epitope; non-shed extracellular portion;
KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
XX breast cancer; ovarian cancer.

OS Homo sapiens.

XX WO2004005470-A2.

XX PD 15-JAN-2004.

XX PF 03-JUL-2003; 2003WO-US020907.

XX PR 03-JUL-2002; 2002US-0393094P.

XX (IMMU-) IMMUNOGEN INC.

XX Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
PI Water CA;
XX WPI; 2004-091350/09.

PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
PT treating or monitoring malignancies, such as breast or ovarian cancer.
XX Claim 14; SEQ ID NO 16; 113pp; English.

CC This invention relates to a novel isolated monoclonal antibody that
CC specifically binds to an epitope of a non-shed extracellular portion of a
CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
CC useful for the development of compounds with a cytostatic or for gene
CC therapy. The composition and methods are useful in diagnosing, treating
CC or monitoring malignancies, such as breast or ovarian cancer. The present
CC sequence is that of a peptide which was used for raising antibodies to
CC the extracellular, non-shed region of Muc16 and which is claimed in the
CC specification.

XX Sequence 21 AA;

XX SQ

Query Match	100.0%;	Score 107;	DB 8;	Length 21;
Best Local Similarity	100.0%;	Pred. NO. 2.6e-10;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 FRNSSIKSYFSDCCVSTFRSV 21			
DB	1 FRNSSIKSYFSDCCVSTFRSV 21			
RESULT 2				
ID	ADJ66729 standard; protein; 108 AA.			
AC	ADJ66729;			
DT	06-MAY-2004 (first entry)			
DE	Human Muc16 epitope amino acid sequence.			
XX				
KM	monoclonal antibody; epitope; non-shed extracellular portion;			
KW	shed antigen; human; Muc1; Muc16; cytosolatic; gene therapy; malignancy;			
OS	breast cancer; ovarian cancer.			
XX				
XX	Homo sapiens.			
PN	MO2004005470-A2.			
XX				
PD	15-JAN-2004.			
XX				
PF	03-JUL-2003; 2003WO-US020907.			
XX				
PR	03-JUL-2002; 2002US-0393094P.			
XX				
PA	(IMMU-) IMMUNOGEN INC.			
PI	Payne G, Chittenden T, Goldmacher V, Chun P, Snelder-Mulready K,			
PI	Vater CA;			
DR	WPI; 2004-091350/09.			
XX				
PT	New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,			
PT	treating or monitoring malignancies, such as breast or ovarian cancer.			
PS	Claim 13; SEQ ID NO 2; 113pp; English.			
XX				
CC	This invention relates to a novel isolated monoclonal antibody that			
CC	specifically binds to an epitope of a non-shed extracellular portion of a			
CC	shed antigen or of human Muc1 or Muc16 protein. The invention may be			
CC	useful for the development of compounds with a cytostatic or for gene			
CC	therapy. The composition and methods are useful in diagnosing, treating			
CC	or monitoring malignancies, such as breast or ovarian cancer. The present			
CC	sequence is that of a Muc epitope of the invention.			
XX				
SO	Sequence 108 AA;			
Query Match	100.0%;	Score 107;	DB 8;	Length 108;
Best Local Similarity	100.0%;	Pred. NO. 1.7e-09;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 FRNSSIKSYFSDCCVSTFRSV 21			
DB	18 FRNSSIKSYFSDCCVSTFRSV 38			
RESULT 3				
ID	ADJ66740 standard; protein; 108 AA.			
AC	ADJ66740;			
XX				
DT	06-MAY-2004 (first entry)			
XX				
DE	Human Muc16 GST fusion protein amino acid sequence SeqID13.			

XX	monoclonal antibody; epitope; non-shed extracellular portion;
KW	shed antigen; human; Muc1; Muc16; cytosolic; gene therapy; malignancy;
KW	breast cancer; ovarian cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO2004005470-A2.
XX	
PD	15-JAN-2004.
XX	
PF	03-JUL-2003; 2003WO-US020907.
XX	
PR	03-JUL-2002; 2002US-0393094P.
XX	
PA	(IMMUNO-) IMMUNOGEN INC.
PI	Payne G, Chittenden T, Goldmacher V, Chun P, Snider-Mulready K,
PI	Vater CA;
XX	
DR	WPI; 2004-091350/09.
XX	
PT	New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
XX	treating or monitoring malignancies, such as breast or ovarian cancer.
XX	
PS	Disclosure; SEQ ID NO 13; 113pp; English.
XX	
CC	This invention relates to a novel isolated monoclonal antibody that
CC	specifically binds to an epitope of a non-shed extracellular portion of a
CC	shed antigen or of human Muc1 or Muc16 protein. The invention may be
CC	useful for the development of compounds with a cytosolic or for gene
CC	therapy. The composition and methods are useful in diagnosing, treating
CC	or monitoring malignancies, such as breast or ovarian cancer. The present
CC	sequence is that of the human Muc16 GST fusion protein which was used in
CC	the exemplification of the invention.
XX	
SO	Sequence 108 AA;
XX	
QY	Query Match 100.0%; Score 107; DB 8; Length 108;
	Best Local Similarity 100.0%; Pred. No. 1.7e-09;
	Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	1 FRNSIKSYFSDCOVSTFRSV 21
	18 FRNSIKSYFSDCOVSTFRSV 38
DB	
XX	
AD	ADM30756 standard; protein; 174 AA.
XX	
AC	ADM30756;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Human cancer linked protein referred to as 192292 Segid 4.
XX	
KW	human; cancer-related gene; screening assay; immuno-conjugate;
KW	cytosolic; immunostimulant; cancer; ovarian cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO2003075854-A2.
XX	
PD	18-SEP-2003.
XX	
PF	07-MAR-2003; 2003WO-US007147.
XX	
PR	07-MAR-2002; 2002US-0362527P.
XX	
PA	(AVAL-) AVALON PHARM.
PI	Young PE, Ebner R, Strovel JW;


```
XX WPI; 2003-748320/70.
DR N-PSDB; ADM30755.
XX
PT Identifying an agent that modulates the activity of any of three specific
PT cancer-related genes, potentially useful in treating (ovarian) cancer,
PT comprises detecting a difference in expression of the gene in the
PT presence of the agent.
XX
PS Claim 11; SEQ ID NO 4; 57bp; English.
XX
CC This invention relates to a novel method of identifying an agent that
CC modulates the activity of a cancer-related gene. Specifically, it refers
CC to a screening assay to identify potential anticancer agents, as well as
CC methods to assess the cancerous state of a cell. The present invention
CC describes antibodies against the expression products of these cancer-
CC related genes that are capable of targeting a cancerous cell in vivo.
CC Furthermore it provides immuno-conjugates containing such antibodies,
CC which can be used to deliver target therapeutics thereto. Accordingly,
CC these compositions that exhibit cytostatic and immunostimulant activities
CC are useful for preventing or treating cancer in humans and in particular
CC the treatment of ovarian cancer. This polypeptide sequence is a protein
CC encoded by a human cancer-linked gene of the invention.
XX
SQ Sequence 174 AA;
XX
Query Match          100.0%; Score 107; DB 7; Length 174;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 FRNSSIKSYFSDCQVSTFRSV 21
   |||||
   28 FRNSSIKSYFSDCQVSTFRSV 48
XX
Db
XX
RESULT 5
ABP30979
ID ABP30979 standard; protein; 178 AA.
XX
AC ABP30979;
XX
XX 02-JUL-2002 (first entry)
XX
DE Predicted extracellular domain of 0772P.
XX
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
XX
XX Homo sapiens.
XX
XX WO200206317-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001WO-US022635.
XX
XX 17-JUL-2000; 2000US-00617747.
XX
XX 10-AUG-2000; 2000US-00636801.
XX
XX 20-SEP-2000; 2000US-00667857.
XX
XX 04-APR-2001; 2001US-00827271.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (CORI-) CORIXA CORP.
XX
XX
XX MITCHAM JL, KING GE, ALGATE PA, FLING SP, RETTER MW, FANGER GR;
XX
XX Reed SG, VEDVICK TS, CARTER D, HILL P, ALBONE E;
XX
XX WPI; 2002-164781/21.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
XX protein or its variants, useful for stimulating an immune response in a
XX patient and treating ovarian cancer.
XX
XX Example 13; Page 361-362; 408bp; English.
XX
PS
```

```
XX
XX This invention relates to polypeptides comprising an immunogenic portion
XX of an ovarian carcinoma protein which acts as an immunostimulant and is
XX cytostatic. The polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations and antigen presenting cells that express
XX the polypeptides are useful for stimulating an immune response in a
XX patient and treating ovarian cancer. This sequence represents protein
XX related to the invention
XX
SQ Sequence 178 AA;
XX
Query Match          100.0%; Score 107; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 FRNSSIKSYFSDCQVSTFRSV 21
   |||||
   87 FRNSSIKSYFSDCQVSTFRSV 107
XX
Db
XX
RESULT 6
ADA08642
ID ADA08642 standard; protein; 178 AA.
XX
AC ADA08642;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human 0772P partial protein #12.
XX
XX Human; gene therapy; ovarian cancer; cancer.
XX
XX Homo sapiens.
XX
XX US2003091580-A1.
XX
XX 15-MAY-2003.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (MITC/) MITCHAM J L.
XX
XX (KING/) KING G E.
XX
XX (ALGA/) ALGATE P A.
XX
XX (FLIN/) FLING S P.
XX
XX (RETT/) RETTER M W.
XX
XX (FANG/) FANGER G R.
XX
XX (REED/) REED S G.
XX
XX (VEDV/) VEDVICK T S.
XX
XX (CART/) CARTER D.
XX
XX (HILL/) HILL P.
XX
XX (ALBO/) ALBONE E.
XX
XX
XX MITCHAM JL, KING GE, ALGATE PA, FLING SP, RETTER MW, FANGER GR;
XX
XX Reed SG, VEDVICK TS, CARTER D, HILL P, ALBONE E;
XX
XX WPI; 2003-532352/50.
XX
XX New isolated 0772P polypeptides and polynucleotides, useful in gene
XX therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX cancer.
XX
XX Disclosure; Page 28; 371bp; English.
XX
XX The invention relates to an isolated 0772P polypeptide, which has the
XX structure fully defined in the specification. The composition containing
XX the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX or antigen presenting cells are useful for stimulating an immune response
XX and treating ovarian cancer. Detecting the presence of the
XX polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX carcinoma cDNAs and protein cDNAs were identified using microarray
XX technology. The present sequence represents a human ovarian carcinoma
XX
```

CC antigen.
 XX
 SQ Sequence 178 AA;
 Query Match 100.0%; Score 107; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FRNSIKSYFSDCOVSTFRSV 21
 |||||
 DB 87 FRNSIKSYFSDCOVSTFRSV 107
 RESULT 7
 ADF08985 ADF08985 standard; protein; 178 AA.
 XX
 AC ADF08985;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Secreted ovarian carcinoma antigen seqid 489.
 XX
 KW gene therapy; protein therapy; vaccine; antibody inhibition;
 KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
 KW secreted ovarian carcinoma antigen.
 XX
 OS Homo sapiens.
 XX
 PN US2003124140-A1.
 XX
 PD 03-JUL-2003.
 XX
 PF 17-JUL-2002; 2002US-00198053.
 XX
 PR 17-DEC-1998; 98US-00215681.
 PR 17-DEC-1998; 98US-00216003.
 PR 23-JUN-1999; 99US-00338933.
 PR 24-SEP-1999; 99US-00404879.
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00678857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 PR 17-JUL-2001; 2001US-00907969.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Bangur CS, Retter MW, Fanger GR, Hill P;
 DR WPI; 2003-897152/82.
 XX
 PT Oncogenic nucleic acids useful for the prevention, diagnosis and
 PT treatment of breast cancer.
 XX
 PS Example 13; SEQ ID NO 489; 399pp; English.
 CC The invention describes nucleic acids (I) and the polypeptides (II) they
 CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
 CC treating diseases related to their aberrant expression i.e. breast
 CC cancer. For example, (I) and (II) may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of (II) by expressing
 CC inactive proteins or to supplement the patient's own production of (II).
 CC Additionally, (I) may be used to produce (II), by inserting (I) into a
 CC host cell and culturing the cell to express the protein (II). (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. The host cell may also be used as antigens in the production of
 CC antibodies against (II) and in assays to identify modulators of (II)'s
 CC expression and activity. The anti-(II) antibodies, agonists and
 CC antagonists may be used to regulate expression and activity and as

CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
 CC immunoassay). This sequence represents a secreted ovarian carcinoma
 CC antigen.
 XX
 SQ Sequence 178 AA;
 Query Match 100.0%; Score 107; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FRNSIKSYFSDCOVSTFRSV 21
 |||||
 DB 87 FRNSIKSYFSDCOVSTFRSV 107
 RESULT 8
 ABP30978 ABP30978 standard; protein; 233 AA.
 XX
 AC ABP30978;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Extracellular and transmembranal regions of 0772P.
 XX
 KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200206317-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-US022635.
 XX
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00678857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR,
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone B;
 XX
 DR WPI; 2002-164781/21.
 XX
 PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 PS Example 13; Page 361; 408pp; English.
 CC This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents protein
 CC related to the invention
 CC
 SQ Sequence 233 AA;
 Query Match 100.0%; Score 107; DB 5; Length 233;
 Best Local Similarity 100.0%; Pred. No. 4.1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FRNSIKSYFSDCOVSTFRSV 21
 |||||
 DB 87 FRNSIKSYFSDCOVSTFRSV 107

```
RESULT 9
ADA08641
ID ADA08641 standard; protein; 233 AA.
XX
XX
XX ADA08641;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human O772P partial protein #11.
XX
XX human; gene therapy; ovarian cancer; cancer.
XX
XX Homo sapiens.
XX
XX US2003091580-A1.
XX
XX 15-MAY-2003.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (MITC/) MITCHAM J L.
XX (KING/) KING G E.
XX (ALGA/) ALGATE P A.
XX (FLIN/) FLING S P.
XX (RETT/) RETTER M W.
XX (FANG/) FANGER G R.
XX (REED/) REED S G.
XX (VEDV/) VEDVICK T S.
XX (CART/) CARTER D.
XX (HILL/) HILL P.
XX (ALBO/) ALBONE E.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
XX New isolated O772P polypeptides and polynucleotides, useful in gene
XX therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX cancer.
XX
XX Example 13; Page 28; 371pp; English.
XX
XX The invention relates to an isolated O772P polypeptide, which has the
XX structure fully defined in the specification. The composition containing
XX the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX or antigen presenting cells are useful for stimulating an immune response
XX and treating ovarian cancer. Detecting the presence of the
XX polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX carcinoma cDNAs and protein cDNAs were identified using microarray
XX technology. The present sequence represents a human ovarian carcinoma
XX antigen.
XX
XX Sequence 233 AA;
XX
XX Query Match 100.0%; Score 107; DB 7; Length 233;
XX Best Local Similarity 100.0%; Pred. No. 4.1e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FRNSSIKSYFSDQVSTFRSV 21
XX |||||
XX Db 87 FRNSSIKSYFSDQVSTFRSV 107
XX
XX RESULT 10
XX ADF08984
XX ID ADF08984 standard; peptide; 233 AA.
XX
XX ADF08984;
XX
XX 12-FEB-2004 (first entry)
```

```
XX
XX Secreted ovarian carcinoma antigen segid 488.
XX
XX DE gene therapy; protein therapy; vaccine; antibody inhibition;
XX KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
XX KW secreted ovarian carcinoma antigen.
XX
XX OS Homo sapiens.
XX
XX PN US2003124140-A1.
XX
XX 03-JUL-2003.
XX
XX 17-JUL-2002; 2002US-00198053.
XX
XX 17-DEC-1998; 98US-00215681.
XX 17-DEC-1998; 98US-00216003.
XX 23-JUN-1999; 99US-00338933.
XX 24-SEP-1999; 99US-00404879.
XX 17-JUL-2000; 2000US-00617747.
XX 10-AUG-2000; 2000US-00636801.
XX 20-SEP-2000; 2000US-00667857.
XX 04-APR-2001; 2001US-00827271.
XX 18-JUN-2001; 2001US-00884441.
XX 17-JUL-2001; 2001US-00907969.
XX
XX (CORI-) CORIXA CORP.
XX
XX Bangur CS, Retter MW, Fanger GR, Hill P;
XX
XX WPI; 2003-897152/82.
XX
XX Oncogenic nucleic acids useful for the prevention, diagnosis and
XX treatment of breast cancer.
XX
XX Example 13; SEQ ID NO 488; 399pp; English.
XX
XX The invention describes nucleic acids (I) and the polypeptides (II) they
XX encode. The nucleic acids (I) may be used for preventing, diagnosing and
XX treating diseases related to their aberrant expression i.e. breast
XX cancers. For example, (I) and (II) may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of (II) by expressing
XX inactive proteins or to supplement the patients own production of (II).
XX Additionally, (I) may be used to produce (II), by inserting (I) into a
XX host cell and culturing the cell to express the protein (II). (I) And its
XX complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar nucleic acids in
XX samples, and therefore which patients may be in need of restorative
XX therapy. The host cell may also be used as antigens in the production of
XX antibodies against (II) and in assays to identify modulators of (II)'s
XX expression and activity. The anti-(II) antibodies, agonists and as
XX CC antagonists may be used to regulate expression and activity and as
XX CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
XX CC immunoassay). This sequence represents a secreted ovarian carcinoma
XX antigen.
XX
XX Sequence 233 AA;
XX
XX Query Match 100.0%; Score 107; DB 7; Length 233;
XX Best Local Similarity 100.0%; Pred. No. 4.1e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 FRNSSIKSYFSDQVSTFRSV 21
XX |||||
XX Db 87 FRNSSIKSYFSDQVSTFRSV 107
XX
XX RESULT 11
XX ABUS4859
XX ID ABUS4859 standard; protein; 284 AA.
XX
XX ABUS4859;
XX
```

XX 12-MAR-2003 (first entry)
 DT Human CA125 carboxy terminal domain.
 DE
 XX
 KM Human; CA125; protein repeat; chromosome 19q 13.2; amino terminal domain;
 KM amino terminal extension; carboxy terminal domain; vaccine; cancer;
 KM ovarian cancer; carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200283866-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US011734.
 XX
 PR 17-APR-2001; 2001US-0284175P.
 PR 19-JUN-2001; 2001US-0299380P.
 PR 27-SEP-2001; 2001US-00965738.
 PR 21-DEC-2001; 2001US-0345180P.
 XX
 PA (UNIV ARKANSAS.
 XX
 PI O'Brien T, Beard J, Underwood L;
 DR WPI; 2003-093013/08.
 XX
 PT New CA125 molecules, useful as a gold standard for detecting and
 PT monitoring the presence of CA125 antigen which can be used for
 PT diagnosing, monitoring or treating patients with cancer or for developing
 PT vaccine against cancer.
 PS
 XX
 PS Claim 1; Fig 9b; 694pp; English.

XX The invention relates to a CA125 protein comprising: (a) an extracellular
 CC amino terminal domain; (b) an amino terminal extension; (c) a multiple
 CC repeat domain; and (d) a carboxy terminal domain. The extracellular amino
 CC terminal domain comprises 5 genomic exons, the amino terminal extension
 CC comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and
 CC the carboxy terminal domain comprises a transmembrane anchor with a short
 CC cytoplasmic domain, and further comprises 9 genomic exons. The gene for
 CC CA125 is located on human chromosome 19q 13.2. Also included are isolated
 CC CA125 repeat domains, nucleic acids (including variants, homologues and
 CC degenerate versions) encoding CA125 proteins or repeat units, a vector
 CC comprising the nucleic acid, a cultured cell comprising the vector, a
 CC method of expressing CA125 antigen in a cell, the amino acid sequences of
 CC the CA125 repeat units (or their variants, fragments or sequences 50%
 CC identical to them), a purified antibody that selectively binds to an
 CC epitope in the receptor-binding domain of CA125 protein, a diagnostic for
 CC detecting and monitoring the presence of CA125 antigen (comprising
 CC recombinant CA125 having at least one repeat unit of the CA125 repeat
 CC domain including epitope binding sites), a therapeutic vaccine to treat
 CC mammals with elevated CA125 antigen levels or at risk of developing a
 CC disease or disease recurrence associated with elevated CA125 antigen
 CC levels (comprising recombinant CA125 repeat domains including epitope
 CC binding sites) and an antisense oligonucleotide that inhibits the
 CC expression of CA125. The CA125 molecule, particularly the multiple repeat
 CC domains are useful as a gold standard for detecting and monitoring the
 CC presence of CA125 antigen, which can be used for diagnosing, monitoring
 CC or treating patients with ovarian cancer and other carcinomas where CA125
 CC is expressed. The molecules are also useful for developing a vaccine
 CC against cancer. The present sequence is a CA125 repeat protein
 XX
 XX
 90 Sequence 284 AA;

Query Match 100.0%; Score 107; DB 6; Length 284;
 Best Local Similarity 100.0%; Pred. No. 5.1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
 DB 138 FRNSIKSYFSDQVSTFRSV 158

RESULT 12
 ID ABP31025 standard; protein; 318 AA.
 XX
 AC ABP31025;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Amino acid sequence of the 3' constant region of 0772P.
 XX
 KM Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200206317-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-US022635.
 XX
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone B;
 XX
 DR WPI; 2002-164781/21.
 XX
 PS Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 PS Claim 3; Page 399-400; 408pp; English.

XX This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents protein
 CC related to the invention
 XX
 XX
 90 Sequence 318 AA;

Query Match 100.0%; Score 107; DB 5; Length 318;
 Best Local Similarity 100.0%; Pred. No. 5.8e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
 DB 172 FRNSIKSYFSDQVSTFRSV 192

RESULT 13
 ID ADA08747 standard; protein; 318 AA.
 XX
 AC ADA08747;
 XX
 PD 06-NOV-2003 (first entry)

DT Human O772P constant region consensus sequence #1.
 DE
 XX
 KM human; gene therapy; ovarian cancer; cancer.
 XX
 OS Homo sapiens.

QY 1 FRNSIKSYFSDQVSTFRSV 21
 DB 138 FRNSIKSYFSDQVSTFRSV 158

```

XX XX US2003091580-A1.
XX PN
XX 15-MAY-2003.
XX PD
XX 17-JUL-2001; 2001US-00907969.
XX PF
XX 18-JUN-2001; 2001US-00844441.
XX PR
XX (MITC/) MITCHAM J L.
XX (KING/) KING G E.
XX (ALGA/) ALGATE P A.
XX (FLIN/) FLING S P.
XX (RETT/) RETTER M W.
XX (FANG/) FANGER G R.
XX (REED/) REED S G.
XX (VEDV/) VEDVICK T S.
XX (CART/) CARTER D.
XX (HILL/) HILL P.
XX (ALBO/) ALBONE E.
XX PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX DR WPI; 2003-532352/50.
XX PT New isolated 0772p polypeptides and polynucleotides, useful in gene
XX PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX PT cancer.
XX PS Claim 3; SEQ ID NO 594; 371pp; English.
XX CC The invention relates to an isolated 0772p polypeptide, which has the
XX CC structure fully defined in the specification. The composition containing
XX CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX CC or antigen presenting cells are useful for stimulating an immune response
XX CC and treating ovarian cancer. Detecting the presence of the
XX CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX CC carcinoma cDNAs and protein cDNAs were identified using microarray
XX CC technology. The present sequence represents a human ovarian carcinoma
XX CC antigen.
XX CC
XX SQ Sequence 318 AA;
XX
XX Query Match 100.0%; Score 107; DB 7; Length 318;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 FRNSIKSYFSDQVSTFRSV 21
XX |||||
XX 172 FRNSIKSYFSDQVSTFRSV 192
XX Db
XX
XX RESULT 14
XX ADF09090
XX ID ADF09090 standard; protein; 318 AA.
XX AC ADF09090;
XX XX
XX DT 12-FEB-2004 (first entry)
XX XX
XX DE Secreted ovarian carcinoma antigen seqid 594.
XX KW gene therapy; protein therapy; vaccine; antibody inhibition;
XX KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
XX KW secreted ovarian carcinoma antigen.
XX OS Homo sapiens.
XX XX
XX PN US2003124140-A1.
XX PD 03-JUL-2003.
XX XX

```

```

PF 17-JUL-2002; 2002US-00198053.
XX XX
XX 17-DEC-1998; 98US-00215681.
XX 17-DEC-1998; 98US-00216003.
XX 23-JUN-1999; 99US-00338933.
XX 24-SEP-1999; 99US-00404879.
XX 17-JUL-2000; 2000US-00617747.
XX 10-AUG-2000; 2000US-00636801.
XX 20-SEP-2000; 2000US-00667857.
XX 04-APR-2001; 2001US-00827271.
XX 18-JUN-2001; 2001US-00844441.
XX 17-JUL-2001; 2001US-00907969.
XX XX
XX (CORI-) CORIXA CORP.
XX PA
XX PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX DR WPI; 2003-697152/82.
XX DR N-PSDB; ADF09064.
XX XX
XX PT Oncogenic nucleic acids useful for the prevention, diagnosis and
XX PT treatment of breast cancer.
XX PS Example 16; SEQ ID NO 594; 399pp; English.
XX XX
XX CC The invention describes nucleic acids (I) and the polypeptides (II) they
XX CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
XX CC treating diseases related to their aberrant expression i.e. breast
XX CC cancers. For example, (I) and (II) may be used to treat disorders
XX CC associated with decreased expression by rectifying mutations or deletions
XX CC in a patient's genome that affect the activity of (II) by expressing
XX CC inactive proteins or to supplement the patient's own production of (II).
XX CC Additionally, (II) may be used to produce (II), by inserting (I) into a
XX CC host cell and culturing the cell to express the protein (II). (I) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acids in
XX CC samples, and therefore which patients may be in need of restorative
XX CC therapy. The host cell may also be used as antigens in the production of
XX CC antibodies against (II) and in assays to identify modulators of (II)'s
XX CC expression and activity. The anti-(II) antibodies, agonists and
XX CC antagonists may be used to regulate expression and activity and as
XX CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
XX CC immunoassay). This sequence represents a secreted ovarian carcinoma
XX CC antigen.
XX CC
XX SQ Sequence 318 AA;
XX
XX Query Match 100.0%; Score 107; DB 7; Length 318;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 FRNSIKSYFSDQVSTFRSV 21
XX |||||
XX 172 FRNSIKSYFSDQVSTFRSV 192
XX Db
XX
XX RESULT 15
XX AAE12631
XX ID AAE12631 standard; protein; 367 AA.
XX AC AAE12631;
XX XX
XX DT 03-JUN-2002 (first entry)
XX XX
XX DE Human gene 1 encoded secreted protein fragment, SEQ ID NO: 27.
XX KW Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;
XX KW respiratory system disorder; asthma; haematopoietic disorder; skin aging;
XX KW immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
XX KW rheumatoid arthritis; inflammation; neurological disorder; septic shock;
XX KW Alzheimer's disease; Parkinson's disease; diabetes; angiogenic disorder;
XX KW atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
XX KW epithelial cell proliferation; transplantation; chemotaxis; infection;

```

KW food additive; wound healing; endocrine disorder; kidney disorder;
 KW gene therapy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200170804-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 16-MAR-2001; 2001WO-US008585.
 XX
 PR 17-MAR-2000; 2000US-0190076P.
 PR 23-AUG-2000; 2000US-0227009P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI, 2001-639119/73.
 XX
 PT Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
 PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
 PT metastases.
 XX
 PS Disclosure; Page 9; 427pp; English.
 XX
 CC AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
 CC ovarian cancer associated protein (collectively known as ovarian cancer
 CC antigens) genes, and AAF12623-AAF12629 represent the proteins they
 CC encode. AAF12630-AAF12638 represent human ovarian cancer antigenic
 CC fragments or variants. Ovarian cancer antigens and their corresponding
 CC DNAs are used in the prevention, diagnosis and treatment of diseases
 CC associated with their inappropriate expression. These disorders include
 CC proliferative disorders, cancer, tumours, respiratory system disorders,
 CC asthma, haematopoietic disorders, diseases of the immune system, AIDS,
 CC skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
 CC inflammation, allergies, neurological disorders (e.g., Alzheimer's
 CC disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, endocrine disorders and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, to identify their cognate ligands or binding
 CC partners, in chemotaxis and can be used as a food additive. Antibodies
 CC specific for a protein of the invention can be used in alleviating
 CC symptoms associated with the disorders mentioned above and in diagnostic
 CC immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
 CC the invention is used in gene therapy. The present sequence represents a
 CC human ovarian cancer antigenic fragment of the invention
 XX
 SQ Sequence 367 AA;
 XX
 Query Match 100.0%; Score 107; DB 4; Length 367;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FRNSSIKSYFSDQVSTFRSV 21
 |||||
 DB 292 FRNSSIKSYFSDQVSTFRSV 312

Search completed: October 27, 2005, 18:44:24
 Job time : 72.6698 secs

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 390
LENGTH: 438
TYPE: PRT
ORGANISM: Homo sapiens
US-09-667-857-390

Query Match 100.0%; Score 107; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 7,3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
|||||
Db 292 FRNSIKSYFSDCOVSTFRSV 312

RESULT 3
US-09-404-879A-389
Sequence 389, Application US/09404879A
Patent No. 6468546

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-09-404-879A-389

Query Match 100.0%; Score 107; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 1,5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
|||||
Db 687 FRNSIKSYFSDCOVSTFRSV 707

RESULT 4
US-09-667-857-389
Sequence 389, Application US/09667857
Patent No. 6699664

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C5
CURRENT APPLICATION NUMBER: US/09/667,857
CURRENT FILING DATE: 2000-09-20
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-09-667-857-389

Query Match 100.0%; Score 107; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 1,5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
|||||
Db 687 FRNSIKSYFSDCOVSTFRSV 707

RESULT 5
US-09-404-879A-312
Sequence 312, Application US/09404879A
Patent No. 6468546

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 312
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapien
US-09-404-879A-312

Query Match 100.0%; Score 107; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1,7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
|||||
Db 768 FRNSIKSYFSDCOVSTFRSV 788

RESULT 6
US-09-338-933-312
Sequence 312, Application US/09338933
Patent No. 6488931

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer Lynn
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
TITLE OF INVENTION: OVARIAN CANCER
FILE REFERENCE: 210121.462C1
CURRENT APPLICATION NUMBER: US/09/338,933
CURRENT FILING DATE: 1999-06-23
NUMBER OF SEQ ID NOS: 312
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 312
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapien
US-09-338-933-312

Query Match 100.0%; Score 107; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1,7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
|||||
Db 768 FRNSIKSYFSDCOVSTFRSV 788

RESULT 7
US-09-667-857-312
Sequence 312, Application US/09667857
Patent No. 6699664

GENERAL INFORMATION:


```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon B.
; APPLICANT: Algate, Paul A.
; APPLICANT: Rling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-667-857-312

Query Match          100.0%; Score 107; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FRNSIKSYFSDQVSTFRSV 21
DB      768 FRNSIKSYFSDQVSTFRSV 788

RESULT 8
US-09-107-433-3451
; Sequence 3451, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <unknown>
; SOFTWARE: <unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3451:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...144
; SEQUENCE DESCRIPTION: SEQ ID NO: 3451:
US-09-107-433-3451

Query Match          43.0%; Score 46; DB 4; Length 144;
Best Local Similarity 56.2%; Pred. No. 3.8;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      3 NSSIKSYFSDQVSTF 18
DB      55 NSSPVSYMRCSLSLF 70

RESULT 9
US-09-425-383-2
; Sequence 2, Application US/09425383
; Patent No. 6194637
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinrui
; TITLE OF INVENTION: Maize DNA Ligase I Orthologue and Uses
; FILE REFERENCE: 0962
; CURRENT APPLICATION NUMBER: US/09/425,383
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/108,793
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 909
; TYPE: PRT
; ORGANISM: Zea mays
US-09-425-383-2

Query Match          43.0%; Score 46; DB 3; Length 909;
Best Local Similarity 33.3%; Pred. No. 32;
Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 FRNSIKSYFSDQVSTF 18
DB      603 FRKATVKSFLDCEIVAY 620

RESULT 10
US-09-787-069-2
; Sequence 2, Application US/09787069
; Patent No. 6627429
; GENERAL INFORMATION:
; APPLICANT: Danisco A/S
; APPLICANT: Christensen, Tove MIE
; APPLICANT: Pedersen, Anette A
; APPLICANT: Brunstedt, Janne
; APPLICANT: Mikkelsen, Jorn D
; TITLE OF INVENTION: Process
; FILE REFERENCE: P005380MO CTH
; CURRENT APPLICATION NUMBER: US/09/787,069
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: GB 9820195.7
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Erwinia chrysanthemi
US-09-787-069-2
```

Query Match 39.3%; Score 42; DB 4; Length 366;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 SIKSYFSDCCQVSTR 16
DB 186 RSFSDCRIS 195

RESULT 11
US-09-710-279-1356
; Sequence 1356, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1356
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1356

Query Match 39.3%; Score 42; DB 4; Length 374;
Best Local Similarity 37.5%; Pred. No. 54;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 SIKSYFSDCCQVSTR 19
DB 132 SEVQNYFKDKQIETVQ 147

RESULT 12
US-09-134-001C-5234
; Sequence 5234, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5234
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5234

Query Match 39.3%; Score 42; DB 3; Length 391;
Best Local Similarity 37.5%; Pred. No. 57;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 SIKSYFSDCCQVSTR 19
DB 149 SEVQNYFKDKQIETVQ 164

RESULT 13
US-09-248-796A-15403
; Sequence 15403, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15403
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15403

Query Match 39.3%; Score 42; DB 4; Length 414;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 SIKSYFSDCCQVSTRSV 21
DB 281 AEVQYFSDALSDGKSV 298

RESULT 14
US-09-442-100-16
; Sequence 16, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mastrock, S. Leelie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:

LENGTH: 561 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-442-100-16

Query Match 39.3%; Score 42; DB 3; length 561;
Best Local Similarity 42.9%; Pred. No. 87;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 FRNSIKSYFSDCQVSTFRSV 21
DB 457 FEHVKMSYFADINFSTLRSM 477

RESULT 15

US-08-939-106-16
Sequence 16, Application US/08939106
Patent No. 6559285
GENERAL INFORMATION:

APPLICANT: Yale University
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATs
GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/939,106
FILING DATE: 26-Nov-6559285-1997
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-007-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 amino acids
TYPE: amino acid

TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-08-939-106-16

Query Match 39.3%; Score 42; DB 4; length 561;
Best Local Similarity 42.9%; Pred. No. 87;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 FRNSIKSYFSDCQVSTFRSV 21
DB 457 FEHVKMSYFADINFSTLRSM 477

Search completed: October 27, 2005, 17:34:33
Job time : 19.2651 secs

THE UNIVERSITY OF CHICAGO

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:30:08 ; Search time 67.2 Seconds
(without alignments)
130.638 Million cell updates/sec

Title: US-10-612-090-16
Perfect score: 107
Sequence: 1 FRNSSIKSYFSDCOVSTFRSV 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 0%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUBCOMB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	21	US-10-612-090-16	Sequence 16, Appl
2	107	100.0	108	US-10-612-090-2	Sequence 2, Appl
3	107	100.0	108	US-10-612-090-13	Sequence 13, Appl
4	107	100.0	174	US-10-383-368-4	Sequence 4, Appl
5	107	100.0	178	US-09-884-441-489	Sequence 489, App
6	107	100.0	178	US-09-907-969-489	Sequence 489, App
7	107	100.0	178	US-10-198-053-489	Sequence 489, App
8	107	100.0	178	US-10-860-790-489	Sequence 489, App
9	107	100.0	233	US-09-884-441-488	Sequence 488, App
10	107	100.0	233	US-09-907-969-488	Sequence 488, App
11	107	100.0	233	US-10-198-053-488	Sequence 488, App

12	107	100.0	233	17	US-10-860-790-488	Sequence 488, App
13	107	100.0	284	10	US-09-965-738-300	Sequence 300, App
14	107	100.0	318	10	US-09-907-969-594	Sequence 594, App
15	107	100.0	318	14	US-10-198-053-594	Sequence 594, App
16	107	100.0	318	17	US-10-860-790-594	Sequence 594, App
17	107	100.0	367	15	US-10-333-900-27	Sequence 27, Appl
18	107	100.0	396	18	US-10-858-412-225	Sequence 225, App
19	107	100.0	438	9	US-09-884-441-390	Sequence 390, App
20	107	100.0	438	9	US-09-884-441-483	Sequence 483, App
21	107	100.0	438	10	US-09-907-969-390	Sequence 390, App
22	107	100.0	438	10	US-09-907-969-483	Sequence 483, App
23	107	100.0	438	10	US-09-827-271-390	Sequence 390, App
24	107	100.0	438	14	US-10-198-053-390	Sequence 390, App
25	107	100.0	438	14	US-10-198-053-483	Sequence 483, App
26	107	100.0	438	15	US-10-333-900-19	Sequence 19, Appl
27	107	100.0	438	17	US-10-860-790-390	Sequence 390, App
28	107	100.0	438	17	US-10-860-790-483	Sequence 483, App
29	107	100.0	439	10	US-09-965-738-148	Sequence 148, App
30	107	100.0	456	18	US-10-858-412-226	Sequence 226, App
31	107	100.0	526	15	US-10-333-900-30	Sequence 30, Appl
32	107	100.0	545	14	US-10-243-243A-4	Sequence 4, Appl
33	107	100.0	583	14	US-10-142-515-4	Sequence 4, Appl
34	107	100.0	748	17	US-10-687-035-1	Sequence 1, Appl
35	107	100.0	809	17	US-10-687-035-2	Sequence 2, Appl
36	107	100.0	833	9	US-09-884-441-389	Sequence 389, App
37	107	100.0	833	10	US-09-907-969-389	Sequence 389, App
38	107	100.0	833	10	US-09-827-271-389	Sequence 389, App
39	107	100.0	833	14	US-10-198-053-389	Sequence 389, App
40	107	100.0	833	17	US-10-860-790-389	Sequence 389, App
41	107	100.0	914	9	US-09-910-689-206	Sequence 206, App
42	107	100.0	914	9	US-09-910-689-206	Sequence 206, App
43	107	100.0	914	9	US-09-884-441-312	Sequence 312, App
44	107	100.0	914	9	US-09-884-441-478	Sequence 478, App
45	107	100.0	914	10	US-09-907-969-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-10-612-090-16
; Sequence 16, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-16

Query Match 100.0%; Score 107; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 FRNSSIKSYFSDCOVSTFRSV 21
DB 1 FRNSSIKSYFSDCOVSTFRSV 21

RESULT 2
US-10-612-090-2
; Sequence 2, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:

```

; APPLICANT: ImmunoGen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-2

Query Match          100.0%; Score 107; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCQVSTFRSV 21
Db 18 FRNSIKSYFSDCQVSTFRSV 38

RESULT 3
US-10-612-090-13
; Sequence 13, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: ImmunoGen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(1)
; OTHER INFORMATION: Glutathione S-transferase fusion site
US-10-612-090-13

Query Match          100.0%; Score 107; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCQVSTFRSV 21
Db 18 FRNSIKSYFSDCQVSTFRSV 38

RESULT 4
US-10-383-368-4
; Sequence 4, Application US/10383368
; Publication No. US20040002091A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul E.
; APPLICANT: Ebner, Reinhard
; APPLICANT: Strovel, Jeffrey W.
; TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
; FILE REFERENCE: 689290-123
; CURRENT APPLICATION NUMBER: US/10/383,368
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US 60/362,527
; PRIOR FILING DATE: 2002-03-07
```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-383-368-4

Query Match          100.0%; Score 107; DB 15; Length 174;
Best Local Similarity 100.0%; Pred. No. 1,2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCQVSTFRSV 21
Db 28 FRNSIKSYFSDCQVSTFRSV 48

RESULT 5
US-09-884-441-489
; Sequence 489, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Derrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-489

Query Match          100.0%; Score 107; DB 9; Length 178;
Best Local Similarity 100.0%; Pred. No. 1,2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCQVSTFRSV 21
Db 87 FRNSIKSYFSDCQVSTFRSV 107

RESULT 6
US-09-907-969-489
; Sequence 489, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-907-969-489

Query Match 100.0%; Score 107; DB 10; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
Db 87 FRNSIKSYFSDQVSTFRSV 107

RESULT 7

US-10-198-053-489
; Sequence 489, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-489

Query Match 100.0%; Score 107; DB 14; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
Db 87 FRNSIKSYFSDQVSTFRSV 107

RESULT 8
US-10-860-790-489
; Sequence 489, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-790-489

Query Match 100.0%; Score 107; DB 17; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
Db 87 FRNSIKSYFSDQVSTFRSV 107

RESULT 9

US-09-884-441-488
; Sequence 488, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-488

Query Match 100.0%; Score 107; DB 9; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
Db 87 FRNSIKSYFSDQVSTFRSV 107

RESULT 10
US-09-907-969-488
; Sequence 488, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-969-488

Query Match 100.0%; Score 107; DB 10; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDQVSTFRSV 21
Db 87 FRNSIKSYFSDQVSTFRSV 107

RESULT 11
US-10-198-053-488
; Sequence 488, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-965-738-300

Query Match          100.0%; Score 107; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ FILE REFERENCE: 210121.462C9
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 488
/ LENGTH: 233
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-198-053-488

RESULT 12
US-10-860-790-488
/ Sequence 488, Application US/10860790
/ Publication No. US20050031634A1
/ GENERAL INFORMATION:
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Hill, Paul
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.462C11
/ CURRENT FILING DATE: 2004-06-02
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 488
/ LENGTH: 233
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-860-790-488

Query Match          100.0%; Score 107; DB 17; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ FILE REFERENCE: 210121.462C9
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 488
/ LENGTH: 233
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-860-790-488

RESULT 13
US-09-965-738-300
/ Sequence 300, Application US/09965738
/ Publication No. US20030143667A1
/ GENERAL INFORMATION:
/ APPLICANT: O'Brien, Timothy
/ TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic
/ FILE REFERENCE: 40715-258841
/ CURRENT FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US 60/284,175
/ NUMBER OF SEQ ID NOS: 306
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 300
/ LENGTH: 284
```

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-965-738-300

Query Match          100.0%; Score 107; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ FILE REFERENCE: 210121.462C9
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 594
/ LENGTH: 318

RESULT 14
US-09-907-969-594
/ Sequence 594, Application US/09907969
/ Publication No. US20030091580A1
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Derrick
/ APPLICANT: Hill, Paul
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.462C8
/ CURRENT FILING DATE: 2001-07-17
/ NUMBER OF SEQ ID NOS: 596
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 594
/ LENGTH: 318
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: variant
/ LOCATION: 136,248,268
/ OTHER INFORMATION: Xaa = Any amino acid
US-09-907-969-594

Query Match          100.0%; Score 107; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ FILE REFERENCE: 210121.462C9
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 594
/ LENGTH: 318

RESULT 15
US-10-198-053-594
/ Sequence 594, Application US/10198053
/ Publication No. US20030124140A1
/ GENERAL INFORMATION:
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Hill, Paul
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.462C9
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 594
/ LENGTH: 318
```


TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: 136,248,268
OTHER INFORMATION: Xaa = Any amino acid
US-10-198-053-594

Query Match 100.0%; Score 107; DB 14; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FRNSSIKSYFSDQCVSTFRSV 21
Db 172 FRNSSIKSYFSDQCVSTFRSV 192

Search completed: October 27, 2005, 18:31:46
Job time : 67.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 18:13:49 ; Search time 12.9907 Seconds
(without alignments)
155.538 Million cell updates/sec

Title: US-10-612-090-16
Perfect score: 107
Sequence: 1 FRNSSIKSYFSDQVSTFRSV 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	54	50.5	496	2	E71192	probable TRK syste
2	49	45.8	493	2	G75151	trk potassium uptr
3	48	44.9	809	2	T40574	guanine nucleotide
4	47	43.9	902	2	S33918	dynammin-like prote
5	44	41.1	51	2	S42357	probable zinc fing
6	44	41.1	145	2	C70113	probable transcrip
7	43	40.2	361	2	AG0052	probable pectinest
8	43	40.2	401	2	F97260	uncharacterized co
9	43	40.2	790	2	S71278	DNA ligase (ATP) (
10	43	40.2	1148	2	T00016	minor outer capsid
11	42	39.3	366	2	S03770	pectinesterase (EC
12	42	39.3	366	2	JN0799	pectinesterase (EC
13	42	39.3	368	2	E87753	protein C43E11.5 (
14	42	39.3	368	2	T29779	hypothetical prote
15	42	39.3	494	2	AB0252	glucose-6-phosphat
16	42	39.3	572	2	S64387	protein kinase DBF
17	41.5	38.8	298	2	F70140	glycerol-3-phospha
18	41	38.3	179	2	F75392	hypothetical prote
19	41	38.3	414	2	C86342	F9H16.7 protein-
20	41	38.3	903	2	T20804	hypothetical prote
21	41	38.3	972	2	F71608	hypothetical prote
22	41	38.3	1520	2	B82274	conserved hypotet
23	41	38.3	1612	2	JC5210	DNA (cytosine-5-)-
24	41	38.3	2632	2	T18718	dynamin heavy chain
25	40.5	37.9	93	2	T51188	small zinc finger-
26	40.5	37.9	310	2	S20889	superantigen Mtv1
27	40.5	37.9	312	2	S34635	superantigen Mtv1
28	40.5	37.9	315	2	JH0554	superantigen Mtv13
29	40.5	37.9	315	2	JH0551	superantigen Mtv1/

30	40.5	37.9	320	2	S26174	superantigen Mtv17
31	40.5	37.9	322	2	S24574	superantigen Mtv15
32	40.5	37.9	2014	2	S46622	probable membrane
33	40	37.4	115	2	T33139	hypothetical prote
34	40	37.4	172	2	AD0062	conserved hypotet
35	40	37.4	236	2	S55297	hypothetical prote
36	40	37.4	342	2	E97522	sensory rhodopsin
37	40	37.4	342	2	E97522	peptide chain rele
38	40	37.4	362	2	AF2741	hypothetical prote
39	40	37.4	382	2	A38545	hypothetical prote
40	40	37.4	382	2	E97742	hypothetical prote
41	40	37.4	429	2	T29711	hypothetical prote
42	40	37.4	433	2	G90083	hypothetical prote
43	40	37.4	444	2	T15907	hypothetical prote
44	40	37.4	468	2	S47447	hypothetical prote
45	40	37.4	491	1	D64947	glucose-6-phosphat

ALIGNMENTS

RESULT 1
E71192
probable TRK system potassium uptake protein - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: E71192
M.: Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusihida, N.; Oguchi, N.; Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic at
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71192
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-496 <KAW>
A:Cross-references: UNIPROT:O59477; GB:AP000007; NID:g3236134; PIDN:BA30932.1; PID:g3256
A:Experimental source: strain O73
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1813

Query Match
Best Local Similarity 50.5%; Score 54; DB 2; Length 496;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRNSSIKSYFSDQVST 17
Db 270 FKNRSIRSPFMDIQVT 286

RESULT 2
G75151
trk potassium uptake system protein (trkh) PAB0248 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: G75151
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A:Reference number: A75001
A:Accession: G75151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <KAW>
A:Cross-references: UNIPROT:Q9V1Q5; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49294
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: trkh, PAB0248
C:Superfamily: potassium uptake protein trkh

Query Match
Best Local Similarity 45.8%; Score 49; DB 2; Length 493;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRNSSIKSYFSDCCQVST 17
 ||:|||||
 DB 266 FRDKSLRNFNDIOYKT 282

RESULT 3
 T40574
 guanine nucleotide binding protein beta subunit-like - fission yeast (Schizosaccharomyces
 C.Species: Schizosaccharomyces pombe
 C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 A.Accession: T40574
 R.Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999
 A.Reference number: Z21938
 A.Accession: T40574
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-809 <SBE>
 A.Cross-references: UNIPROT:Q94527; EMBL:AL035226; PIDN:CAA22832.1; GSPDB:GN00067
 A.Experimental source: strain 972h-; cosmid c609
 C.Genetics:
 A.Gene: SPAC609.03
 A.Map position: 2
 A.Introns: 21/1; 52/2; 280/3; 780/2

Query Match 44.9%; Score 48; DB 2; Length 809;
 Best Local Similarity 36.8%; Pred. No. 12;
 Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 NSSIKSYFSDCCQVSTFRSY 21
 ||:|||||
 DB 658 NTHVKSYYGHCNVEISIKNV 676

RESULT 4
 S33918
 dynamn-like protein MGM1 - yeast (Saccharomyces cerevisiae)
 N.Alternate names: protein Os010; protein YOR211c; protein YOR50-1
 C.Species: Saccharomyces cerevisiae
 C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 A.Accession: S33918; S36728; A42137; S48883; S6038; S6T104; S6T103; S71713
 R.Guan, K.; Fath, U.; Marshall, T.K.; Deschenes, R.U.
 Curr. Genet. 24, 141-148, 1993
 A>Title: Normal mitochondrial structure and genome maintenance in yeast requires the dyn
 A.Reference number: S33918; MUID:93365024; PMID:7916673
 A.Accession: S33918
 A.Molecule type: DNA
 A.Residues: 1-902 <GUA1>
 A.Cross-references: UNIPROT:P32266; EMBL:L07419
 A.Note: It is uncertain whether Met-1 or Met-60 is the initiator
 R.Guan, K.; Fath, U.; Marshall, T.; Deschenes, R.U.
 submitted to the EMBL Data Library, January 1992
 A.Reference number: S36728
 A.Accession: S36728
 A.Molecule type: DNA
 A.Residues: 1-190, 'A', 192-902 <GUA2>
 A.Cross-references: EMBL:L07419
 R.Jones, B.A.; Pandman, W.L.
 Genes Dev. 6, 380-389, 1992
 A>Title: Mitochondrial DNA maintenance in yeast requires a protein containing a region r
 A.Reference number: A42137; MUID:92192451; PMID:1532158
 A.Accession: A42137
 A.Molecule type: DNA
 A.Residues: 60-149, 'C', 151-902 <JON>
 A.Cross-references: EMBL:X62834; NID:g93956; PIDN:CAA4637.1; PID:g3957
 A.Note: Sequence extracted from NCBI backbone (NCBIN:88065, NCBI:P:88066)
 R.Lialo, D.; Carles, C.; Sentenac, A.; Thuriaux, P.
 Proc. Natl. Acad. Sci. U.S.A. 90, 5524-5528, 1993
 A>Title: Interactions between three common subunits of yeast RNA polymerases I and III.
 A.Reference number: S48883; MUID:93396170; PMID:8516295
 A.Accession: S48883

A>Status: nucleic acid sequence not shown; translation not shown
 A.Molecule type: DNA
 A.Residues: 734-902 <IAL>
 A.Cross-references: EMBL:L11274; NID:g295632; PIDN:AAB59316.1; PID:g295633
 A.Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993
 R.Gallison, F.; Dujon, B.
 submitted to the EMBL Data Library, October 1995
 A.Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
 A.Reference number: S60938
 A.Accession: S60938
 A.Molecule type: DNA
 A.Residues: 1-805 <GAL>
 A.Cross-references: EMBL:X92441; NID:g1050762; PIDN:CAA63174.1; PID:g1050763
 R.Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.
 submitted to the Protein Sequence Database, July 1996
 A.Reference number: S67104
 A.Accession: S67104
 A.Molecule type: DNA
 A.Residues: 1-805 <BOY>
 A.Cross-references: EMBL:Z75119; MIPS:YOR211c
 R.Hughes, B.; Pohl, T.M.
 submitted to the Protein Sequence Database, July 1996
 A.Reference number: S66685
 A.Accession: S67103
 A.Molecule type: DNA
 A.Residues: 179-902 <HUG>
 A.Cross-references: EMBL:Z75119; MIPS:YOR211c
 A.Experimental source: strain S288C
 R.Gallison, F.; Dujon, B.
 yeast 12, 877-885, 1996
 A>Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of
 A.Reference number: S71713; MUID:96437977; PMID:8840505
 A.Accession: S71713
 A.Molecule type: DNA
 A.Residues: 1-150 <GAW>
 A.Cross-references: EMBL:X92441
 C.Genetics:
 A.Gene: SGD:MGM1
 A.Cross-references: SGD:S0005737; MIPS:YOR211c
 A.Map position: 15R
 A.Genome: nuclear
 C.Function:
 A.Description: mitochondrial genome maintenance
 C.Keywords: GTP binding; mitochondrion; nucleotide binding; P-loop
 F/238-245/Region: nucleotide-binding motif A (P-loop)

Query Match 43.9%; Score 47; DB 2; Length 902;
 Best Local Similarity 72.7%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 KSYFSDCCQVST 17
 ||:|||||
 DB 496 KKYFTNCQVST 506

RESULT 5
 S42357
 probable zinc finger protein - common cuttlefish
 C.Species: Sepia officinalis (common cuttlefish)
 C.Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
 A.Accession: S42357
 R.Martin-Poncheu, A.; Mouters-Tyrou, D.; Pudlo, B.; Buissine, E.; Sautiere, P.
 Eur. J. Biochem. 220, 463-468, 1994
 A>Title: Isolation and characterization of a small putative zinc finger protein from cutt
 A.Reference number: S42357; MUID:94170795; PMID:8125104
 A.Accession: S42357
 A.Status: preliminary
 A.Molecule type: protein
 A.Residues: 1-51 <MAR>
 A.Cross-references: UNIPROT:P80304
 C.Keywords: DNA binding; nucleus

RESULT 8

A;Description: involved in vector transmissibility

C:Superfamily: rice dwarf virus minor outer capsid protein, 127K

Query Match 40.2%; Score 43; DB 2; Length 1148;

Best Local Similarity 53.8%; Pred. No. 1,le02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 YFSDCQVSTFRSV 21

Db 326 YITECEVDLTRSV 338

RESULT 11

S03770 pectinesterase (EC 3.1.1.11) precursor - Erwinia chrysanthemi

N:Alternate names: pectin methyl esterase

C:Species: Erwinia chrysanthemi

C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004

C:Accession: S03770

R:Plasmid, G.S.

Mol. Microbiol. 2, 247-254, 1988

A:Title: Molecular cloning and nucleotide sequence of the pectin methyl esterase gene of

A:Reference number: S03770; MUID:88246052; PMID:2837615

A:Accession: S03770

A:Molecule type: DNA

A:Residues: 1-366 <PLA>

A:Cross-references: UNIPROT:P07863; GB:Y00549; NID:G42442; PID:CAA66628.1; PID:G42443

C:Genetics:

A:Gene: pme

C:Superfamily: pectinesterase

C:Keywords: carboxylic ester hydrolase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-366/Product: pectin esterase #status predicted <MAT>

Query Match 39.3%; Score 42; DB 2; Length 366;

Best Local Similarity 60.0%; Pred. No. 51;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 KSYFSDQCVS 16

Db 186 RSFSDCRIS 195

RESULT 12

UN0799 pectinesterase (EC 3.1.1.11) precursor - Erwinia chrysanthemi

N:Alternate names: pectin methyl esterase

C:Species: Erwinia chrysanthemi

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C:Accession: UN0799; FN0623

R:Laurent, F.; Kotoujansky, A.; Labasse, G.; Bertheau, Y.

Gene 131, 17-25, 1993

A:Title: Characterization and overexpression of the pme gene encoding pectin methyl ester

A:Reference number: UN0799; MUID:93380669; PMID:8370537

A:Accession: UN0799

A:Molecule type: DNA

A:Residues: 1-366 <LAU>

A:Cross-references: UNIPROT:P07863; GB:L07644; NID:G148465; PIDN:AAA24852.1; PID:G148466

A:Experimental source: strain 3937

A:Accession: FN0623

A:Molecule type: protein

A:Residues: 25-34 <LAU>

C:Comment: This enzyme is required for the full development of the soft rot disease of r

C:Genetics:

A:Gene: pem

C:Superfamily: pectinesterase

C:Keywords: carboxylic ester hydrolase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-366/Product: pectinesterase #status predicted <MAT>

Query Match 39.3%; Score 42; DB 2; Length 366;

Best Local Similarity 60.0%; Pred. No. 51;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 KSYFSDQCVS 16

Db 186 RSFSDCRIS 195

RESULT 13

B87753 protein C43B11.5 (imported) - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: B87753

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_eleg

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: B87753

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <STO>

A:Cross-references: UNIPROT:P91147; GB:chr_I; PIDN:AA37621.1; PID:G1703567; GSPDB:GN0001

C:Genetics:

A:Gene: C43B11.5

A:Map position: 1

Query Match 39.3%; Score 42; DB 2; Length 366;

Best Local Similarity 53.8%; Pred. No. 51;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRNSISKYSFSDC 13

Db 312 FRNGISITYOFC 324

RESULT 14

T29779

hypothetical protein C50F2.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29779

R:Du, Z.; Le, T.T.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans cosmid C50F2.

A:Reference number: Z20684

A:Accession: T29779

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-368 <DUZ>

A:Cross-references: UNIPROT:P91179; EMBL:U80445; PIDN:AA37790.1; GSPDB:GN00019; CESP:CS

A:Experimental source: strain Bristol N2; clone C50F2

C:Genetics:

A:Gene: CESP:C50F2.5

A:Map position: 1

A:Introns: 27/3; 130/3; 257/3; 342/3

Query Match 39.3%; Score 42; DB 2; Length 366;

Best Local Similarity 53.8%; Pred. No. 51;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRNSISKYSFSDC 13

Db 312 FRNGISITYOFC 324

RESULT 15

AB0252

glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) [imported] - Yersinia pestis (strain CC

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB0252

R:Parthill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Farrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F

11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MIMD:21470413; PMID:11586360
 A/Accession: AB0252
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-494 <KUR>
 A/Cross-references: UNIPROT:Q8ZET7; GB:AL590842; PIDN:CAC90878.1; PID:G15980077; GSPDB:C
 C/Genetics: zwf
 C/Superfamily: glucose-6-phosphate dehydrogenase
 C/Keywords: oxidoreductase

Query Match 39.3%; Score 42; DB 2; Length 494;
 Best Local Similarity 46.2%; Pred. No. 69;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 NSSIKSYFSPDQV 15
 | : ||: ||
 Db 163 NDQVAEYFNCCQV 175

Search completed: October 27, 2005, 18:57:26
 Job time : 13.9907 secs

George Washington

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:34:49 ; Search time 61.0465 Seconds

(without alignments)
176.155 Million cell updates/sec

Title: US-10-612-090-16

Perfect score: 107
Sequence: 1 FRNSSIKSYFSDCVSTFRSV 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	1148	Q9H7S7	Q9H7S7 homo sapien
2	107	100.0	6995	Q26R82	Q26R82 homo sapien
3	107	100.0	22152	Q8W417	Q8W417 homo sapien
4	96	89.7	258	Q9D1H1	Q9D1H1 mus musculus
5	54	50.5	436	Q59477	Q59477 pyrococcus
6	51	47.7	363	Q6CS86	Q6CS86 kluyveriomy
7	51	47.7	663	Q7RF25	Q7RF25 plasmodium
8	51	47.7	955	Q8IDT3	Q8IDT3 plasmodium
9	49	45.8	493	Q9V105	Q9V105 pyrococcus
10	49	45.8	548	Q75103	Q75103 oryza sativ
11	49	45.8	1524	Q7RCM2	Q7RCM2 plasmodium
12	48	44.9	411	Q86128	Q86128 dictyostell
13	48	44.9	809	Q94527	Q94527 schizosacch
14	47	43.9	253	TRPC_BACCR	TRPC_BACCR
15	47	43.9	881	MGM1_YEAST	MGM1_YEAST
16	47	43.9	1235	Q8PF69	Q8PF69 saccharomy
17	46	43.0	253	Q73H09	Q73H09 xanthomonas
18	46	43.0	185	Q8T2W9	Q8T2W9 pyrococcus
19	45	42.1	505	Q8TPM5	Q8TPM5 methanocarc
20	45	42.1	511	Q8TPW7	Q8TPW7 methanocarc
21	45	42.1	511	Q8TRD8	Q8TRD8 methanocarc
22	45	42.1	602	Q9W712	Q9W712 pseudopleur
23	45	42.1	682	Q8EW4	Q8EW4 mycoplasma
24	45	42.1	891	Q6FLK2	Q6FLK2 candida gla
25	45	42.1	1812	Q81538	Q81538 plasmodium
26	44.5	41.6	847	Q93XL9	Q93XL9 rosa hybrid
27	44	41.1	51	SPE_SEPOF	SPE_SEPOF
28	44	41.1	131	Q81BE9	Q81BE9 enterococcu
29	44	41.1	145	NUS8_BORBU	NUS8_BORBU
30	44	41.1	174	Q6B365	Q6B365 uncultured
31	44	41.1	187	Q7PMW5	Q7PMW5 anopheles g

32	44	41.1	270	Q757C9	Q757C9 aabhya goss
33	44	41.1	274	Q6LPV1	Q6LPV1 photobacter
34	44	41.1	330	Q8GC33	Q8GC33 leuconostoc
35	44	41.1	583	Q6MMQ0	Q6MMQ0 bdellovibri
36	44	41.1	810	Q8W5J3	Q8W5J3 oryza sativ
37	44	41.1	810	Q7XD67	Q7XD67 oryza sativ
38	44	41.1	933	Q7MG24	Q7MG24 vibrio vuln
39	44	41.1	933	Q8D4J2	Q8D4J2 vibrio vuln
40	44	41.1	1063	Q934I4	Q934I4 streptococc
41	44	41.1	1063	Q6W9D8	Q6W9D8 streptococc
42	43	40.2	122	Q9NA57	Q9NA57 caenorhabdi
43	43	40.2	212	Q8EBH3	Q8EBH3 drosophila
44	43	40.2	253	TRPC_BACAN	TRPC_BACAN
45	43	40.2	253	Q63EC9	Q63EC9 bacillus ce

ALIGNMENTS

RESULT 1
Q9H7S7 PRELIMINARY; PRT; 1148 AA.
ID Q9H7S7
AC Q9H7S7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; DOI=10.1038/ng1285;
RX PubMed=14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Iehibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimura M., Watanabe M., Hirooka S., Chiba Y., Iehida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamagaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro K., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Moritaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Mesuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL, AK024365; BAB14899.1; -;
DR HSSP; Q9D1H1; 11VZ.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 7.
DR PROSITE; PS50024; SEA; 3.
SQ SEQUENCE 1148 AA; 127957 MW; 3861B0D5EFD8BAC CRC64;

```

Query Match          100.0%; Score 107; DB 2; Length 1148;
Best Local Similarity 100.0%; Pred. No. 5e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCOVSTFRSV 21
    |||||
Db 1002 FRNSIKSYFSDCOVSTFRSV 1022

RESULT 2
O96RK2 ID O96RK2 PRELIMINARY; PRT; 6995 AA.
AC O96RK2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mucin 16 (Fragment).
GN Name=MUC16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN 1
  [1]
  SEQUENCE FROM N.A.
  RX MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200;
  RA Yin B.W., Lloyd K.O.;
  RT "Molecular cloning of the cal25 ovarian cancer antigen. Identification
  as a new mucin, mucin 16."
  RT J. Biol. Chem. 276:27371-27375(2001).
  RN 2
  [2]
  SEQUENCE FROM N.A.
  RP Lloyd K.O., Yin B.W.T.;
  RA Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases.
  DR EMBL; AF361486; AK474120.3; -.
  DR HSSP; Q9D1H1; 11VZ.
  DR InterPro; IPR00194; ATPase_a/bcentre.
  DR InterPro; IPR00082; SEA.
  DR Pfam; PF01390; SEA; 20.
  DR SMART; SM00200; SEA; 10.
  DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
  DR PROSITE; PS00024; SEA; 6.
  RT NON TER
  RT SEQUENCE 6995 AA; 744958 MW; 80C797BDBF33A2B CRC64;

Query Match          100.0%; Score 107; DB 2; Length 6995;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCOVSTFRSV 21
    |||||
Db 6849 FRNSIKSYFSDCOVSTFRSV 6869

RESULT 3
O8WX17 ID O8WX17 PRELIMINARY; PRT; 22152 AA.
AC O8WX17;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Ovarian cancer related tumor marker CA125.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN 1
  [1]
  SEQUENCE FROM N.A.
  RP MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
  RA O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
  RA York L.;
  RT "The CA 125 gene: an extracellular superstructure dominated by repeat
  RT sequences."
  RT Tumour Biol. 22:348-366(2001).

```

```

RN 2
  [2]
  SEQUENCE FROM N.A.
  RP O'Brien T.J., Underwood L.J., Beard J.B.;
  RA Submitted (OCT-2002) to the EMBL/Genbank/DBJ databases.
  DR EMBL; AF144442; AA65133.2; -.
  DR Genew; HGNC:1582; MUC16.
  DR Pfam; PF01390; SEA; 51.
  DR SMART; SM00200; SEA; 23.
  DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_2.
  DR PROSITE; PS00024; SEA; 11.
  RT SEQUENCE 22152 AA; 235268 MW; B3E7BDF19997A440 CRC64;

Query Match          100.0%; Score 107; DB 2; Length 22152;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FRNSIKSYFSDCOVSTFRSV 21
    |||||
Db 22006 FRNSIKSYFSDCOVSTFRSV 22026

RESULT 4
O9D1H1 ID O9D1H1 PRELIMINARY; PRT; 258 AA.
AC O9D1H1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:1110008114 product:hypothetical SEA domain containing
DE protein, full insert sequence.
GN Name=1110008114R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
  [1]
  SEQUENCE FROM N.A.
  RP STRAIN=C57BL/6J; TISSUE=whole body;
  RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
  RA Carninci P., Hayashizaki Y.;
  RT "High-efficiency full-length cDNA cloning."
  RT Meth. Enzymol. 303:19-44(1999).
  RN 2
  [2]
  SEQUENCE FROM N.A.
  RP STRAIN=C57BL/6J; TISSUE=whole body;
  RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
  RA RIKEN FANTOM Consortium;
  RT "Functional annotation of a full-length mouse cDNA collection."
  RT Nature 409:685-690(2001).
  RN 3
  [3]
  SEQUENCE FROM N.A.
  RP STRAIN=C57BL/6J; TISSUE=whole body;
  RA The FANTOM Consortium;
  RT "Analysis of the mouse transcriptome based on functional annotation of
  RT 60,770 full-length cDNAs."
  RT Nature 420:563-573(2002).
  RN 4
  [4]
  SEQUENCE FROM N.A.
  RP STRAIN=C57BL/6J; TISSUE=whole body;
  RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
  RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
  RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
  RT "Normalization and subtraction of cap-trapper-selected cDNAs to
  RT prepare full-length cDNA libraries for rapid discovery of new genes."
  RT Genome Res. 10:1617-1630(2000).
  RN 5
  [5]
  SEQUENCE FROM N.A.
  RP STRAIN=C57BL/6J; TISSUE=whole body;
  RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
  RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
  RA Kono H., Akiyama J., Nishi K., Kitsuai T., Teshiro H., Itoh M.,

```

```
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hazada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.",
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK003577; BAB22869.1; -.
PDB; 1IV2; NMR; A=60-190.
DR MGD; MGI:1920982; 111000811ARik.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 1.
DR PROSITE; PS50024; SEA; 1.
DR KX Hypothetical protein.
SQ SEQUENCE 258 AA; 29425 MW; B64D9B6394D84E7 CRC64;

Query Match 89.7%; Score 96; DB 2; Length 258;
Best Local Similarity 90.5%; Pred. No. 5.8e-07;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVSTFRSV 21
Db 111 FRNSIKSYFSDCOVLAFRSV 131

RESULT 5
ID 059477 PRELIMINARY; PRT; 496 AA.
AC 059477;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 496aa long hypothetical TRK system potassium uptake protein.
GN OrderedocNames=PH1813;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=OT3;
RA MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Hatake Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuuchi Y., Shizuya H., Kikuchi H.,
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.",
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000007; BAA30932.1; -.
DR PIR; E71192; E71192.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008324; F:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR003445; Cat_transpt.
```

```
DR InterPro; IPR004772; K_transpttrk.
DR Pfam; PF02386; TrKH; 1.
DR TrRPFAMs; TrGR00933; 2a38; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 496 AA; 55062 MW; 599A8F1CA8711EA2 CRC64;

Query Match 50.5%; Score 54; DB 2; Length 496;
Best Local Similarity 45.8%; Pred. No. 5.1;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRNSIKSYFSDCOVST 17
Db 270 FKRSIRSFMDIQVKT 286

RESULT 6
ID 06CS86 PRELIMINARY; PRT; 368 AA.
AC 06CS86;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome D of strain NRRL Y-
GN 1140 of Kluyveromyces lactis.
ORFNames=KLIA0D03036g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Duret P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangul L., Aigle M., Anichoud V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boismare A., Boyer J., Cattolico L., Confanier F., de Darvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreya F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Oztier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenene D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Pukhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RT "Genome evolution in yeasts.",
RL Nature 430:35-44(2004).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382124; CAH00299.1; -.
DR InterPro; IPR007852; Cdc73.
DR Pfam; PF05179; Cdc73; 1.
SQ SEQUENCE 368 AA; 42023 MW; 3A6E4A8F6647A108 CRC64;

Query Match 47.7%; Score 51; DB 2; Length 368;
Best Local Similarity 45.0%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 RNSSIKSYFSDCOVSTFRSV 21
Db 72 RNSSIKSYFSDCOVSTFRSV 91

RESULT 7
ID 07RF25 PRELIMINARY; PRT; 663 AA.
AC 07RF25;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
```

DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=PY04556;
 OS Plasmodium yoelii yoelii.
 CC Burkarya; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OK NCBI_TaxID=73239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNL;
 RX PubMed=1236865; DOI=10.1038/nature01099;
 RA Carlton J.M., Angiuoli S.V., Suh B.B., Kocot T.W., Pertea M., Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L., Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L., Shalton S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V., Cho J.K., Quakebush J., Sedegah M., Shoabdi A., Cummings L.M., Florens L., Yates J.R., III, Raine J.D., Sinden R.B., Harris M.A., Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B., van Lin L.H., Janse C.U., Waters A.P., Smith H.O., White O.R., Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J., Carucci D.J.;
 RA "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";
 RT Nature 419:512-519(2002). The sequence shown here is derived from an
 CC -1- CATTON: The sequence shown here is derived from an
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AABL01001396; EAA16437.1; -.
 DR Hypothetical protein.
 KW SEQUENCE 663 AA; 79594 MW; 0A42FCE190F415B0 CRC64;

Query Match 47.7% Score 51; DB 2; Length 663;
 Best Local Similarity 53.3%; Pred. No. 21;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 6 IKSPFSDCOVSTFRS 20
 DB 458 LKRYIECOVPTFRS 472
 ||:|||||:
 ||:|||||:

RESULT 8
 Q81DT3 PRELIMINARY; PRT; 955 AA.
 AC Q81DT3;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein PR13_0219.
 GN Name=PR13_0219;
 OS Plasmodium falciparum (Isolate 3D7).
 CC Burkarya; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OK NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B., Lemard N., Clark L., Line A., Barron A., Corton C., Barrin M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J., Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.;
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL844509; CAD52536.1; -.
 KW Hypothetical protein.
 SO SEQUENCE 955 AA; 114349 MW; 10AB720DBCEFC999 CRC64;

Query Match 47.7% Score 51; DB 2; Length 955;
 Best Local Similarity 41.2%; Pred. No. 30;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 4 SIKSPFSDCOVSTFRS 20
 DB 771 SYFKYVACQIOTFRN 787
 ||:|||||:
 ||:|||||:

RESULT 9
 Q9V105

ID Q9V105 PRELIMINARY; PRT; 493 AA.
 AC Q9V105;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE TrkG/H trk potassium uptake system protein.
 GN ORFNames=PAB0248;
 OS Pyrococcus abyssi.
 CC Pyrococcus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OK NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RX MEDLINE=22511545; PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Gaipier M., Hellig R., Lecompte O., Poch O., Priour D., Querellou J., Rapp R., Thierry J.-C., Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
 RA "An integrated analysis of the genome of the hyperthermophilic archaeon Pyrococcus abyssi.";
 RT Mol. Microbiol. 47:1495-1512(2003).
 RL EMBL; AJ248284; CAB49294.1; -.
 DR PIR; G75151; G75151.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008324; F:cation transporter activity; IEA.
 DR GO; GO:000812; P:cation transport; IEA.
 DR InterPro; IPR003445; Cat. transp.
 DR InterPro; IPR004772; K.transpTrk.
 DR Pfam; PF02386; TrkH; 1.
 DR TIGRFAMs; TIGR00933; 2a38; 1.
 KW Complete proteome.
 SO SEQUENCE 493 AA; 54741 MW; 0B26B7B2B5F81283 CRC64;

Query Match 45.8% Score 49; DB 2; Length 493;
 Best Local Similarity 47.1%; Pred. No. 32;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 FRNSIKSPFSDCOVST 17
 DB 266 FRDKSLRNFRFDIOVKT 282
 ||:|||||:
 ||:|||||:

RESULT 10
 Q75L03 PRELIMINARY; PRT; 548 AA.
 ID Q75L03;
 AC Q75L03;
 DT 05-JUN-2004 (TReMBLrel. 27, Created)
 DT 05-JUN-2004 (TReMBLrel. 27, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE 'contains similar to RNA recognition motif. (A.K.a. RRM, RBD, or RNP domain), PF00076' (Putative RNA recognition motif (RRM)-containing protein).
 GN Name=OSUNB0041A22.21; Synonyms=OJ1654_B10.6;
 OS Oryza sativa (Japonica cultivar-group).
 CC Burkarya; Viridiplantae; Embryophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;
 OC Eubryotidae; Oryzae; Oryza.
 OK NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M., Chao Y.-T., Lee P.-F., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H., Hsiung J.-N., Hsu C.-H., Kau P.-I., Lee M.-C., Liu H.-L., Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W., Wu H.-P., Shaw J.-F.;
 RL Submitted (MAR-2004) to the EMBL/Genbank/DBJ databases.

RESULT 11
 Q9V105 PRELIMINARY; PRT; 493 AA.
 AC Q9V105;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE TrkG/H trk potassium uptake system protein.
 GN ORFNames=PAB0248;
 OS Pyrococcus abyssi.
 CC Pyrococcus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OK NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RX MEDLINE=22511545; PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Gaipier M., Hellig R., Lecompte O., Poch O., Priour D., Querellou J., Rapp R., Thierry J.-C., Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
 RA "An integrated analysis of the genome of the hyperthermophilic archaeon Pyrococcus abyssi.";
 RT Mol. Microbiol. 47:1495-1512(2003).
 RL EMBL; AJ248284; CAB49294.1; -.
 DR PIR; G75151; G75151.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008324; F:cation transporter activity; IEA.
 DR GO; GO:000812; P:cation transport; IEA.
 DR InterPro; IPR003445; Cat. transp.
 DR InterPro; IPR004772; K.transpTrk.
 DR Pfam; PF02386; TrkH; 1.
 DR TIGRFAMs; TIGR00933; 2a38; 1.
 KW Complete proteome.
 SO SEQUENCE 493 AA; 54741 MW; 0B26B7B2B5F81283 CRC64;

RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
Wu H.-P., Shaw J.-F.;
RT "Oryza sativa BAC OJ1654 B10 genomic sequence";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC093921; AA075242.1; -;
DR EMBL: AC108504; AA044075.1; -;
DR GO: GO:0003676; Functional acid binding; IEA.
DR InterPro: IPR006550; A/AMP deam AS.
DR InterPro: IPR000504; RNA rec mot.
DR InterPro: IPR001878; ZnF_CCHC.
DR Pfam: PF00076; RRM_1.2.
DR Pfam: PF00098; ZF_CCHC; 4.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00360; RRM; 2.
DR SMART: SM00343; ZnF_C2HC; 4.
DR PROSITE: PS00485; A_DEAMINASE; UNKNOWN_1.
DR PROSITE: PS0102; RRM; 2.
DR PROSITE: PS0158; ZF_CCHC; 4.
SQ SEQUENCE 548 AA; 60337 MW; 5B71485E26025D75 CRC64;

Query Match 45.8%; Score 49; DB 2; Length 548;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 6 IKSYFSDCQVSTFR 19
Db 283 LKFFSDCKISSIR 296

RESULT 11

Q7RCM2 PRELIMINARY; PRT; 1524 AA.
AC Q7RCM2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
GN Name=PY05757;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=73239;

RA SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=1236885; DOI=10.1038/nature01099;
RA Carlton J.M., Anguioni S.V., Suh B.B., Kooij T.W., Petrea M.,
Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
Choi J.K., Quackenbush J., Sedegah M., Shoshiba A., Cummings L.M.,
Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
Cunningham D.A., Preiser P.R., Bergman L.W., Vaideya A.B.,
van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
Salberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519 (2002).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL: AABL0100186; EAA17831.1; -;
KW EMBL: AABL0100186; EAA17831.1; -;
DR Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1524 AA; 178163 MW; E2E6D84B9EFAFAD CRC64;

Query Match 45.8%; Score 49; DB 2; Length 1524;
Best Local Similarity 44.4%; Pred. No. 1e+02;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNNSSIKSYFSDCQVSTFR 18
Db 1152 FVNDSTRSYLSECFNSF 1169

RESULT 12

Q86128 PRELIMINARY; PRT; 411 AA.
AC Q86128;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to Dictyostelium discoideum (slime mold). MkpA protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
NCBI_TaxID=44689;

RA SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafiranski K., Pachbat J., Dear P.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85 (2002).

Qy 1 FNNSSIKSYFSDCQVSTFRSV 21
Db 229 FNNSSIKSYFSDCQVSTFRSV 247

Query Match 44.9%; Score 48; DB 2; Length 411;
Best Local Similarity 38.1%; Pred. No. 37;
Matches 8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Qy 1 FNNSSIKSYFSDCQVSTFRSV 21
Db 229 FNNSSIKSYFSDCQVSTFRSV 247

RESULT 13

Q94527 PRELIMINARY; PRT; 809 AA.
AC Q94527;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SPB609.03 protein.
GN Name=SPB609.03;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;

RA SEQUENCE FROM N.A.
RC STRAIN=972H-;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,
Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
Woodward J., Wolckardt G., Aert R., Robben J., Grymoprez B.,
Wetjens I., Vanstreels E., Rieger M., Schater M., Muller-Auer S.,
Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,

CC STRAIN=5288c / FY1679;
CC MEDLINE=9643797; PubMed=844505;
CC DOI=10.1002/(SICI)1097-0061(199607)12:9<877::AID-YEA969>3.0.CO;2-S;
CC Galleson F., Dujon B.;
CC "Sequence and analysis of a 33 kb fragment from the right arm of
CC chromosome XV of the yeast *Saccharomyces cerevisiae*.";
CC Yeast 12:877-885(1996).
CC [5]
CC SEQUENCE OF 713-881 FROM N.A.
CC STRAIN=ATCC 28383 / FL100.
CC MEDLINE=93296170; PubMed=8516295;
CC Lalo D., Carles C., Sentenac A., Thuriaux P.;
CC "Interactions between three common subunits of yeast RNA polymerases I
CC and III.";
CC Proc. Natl. Acad. Sci. U.S.A. 90:5524-5528(1993).
CC [6]
CC SEQUENCE OF 60-67 AND 140-147, FUNCTION, AND SUBCELLULAR LOCATION.
CC STRAIN=W303;
CC PubMed=12707284; DOI=10.1074/jbc.M211311200;
CC Herlan M., Vogel F., Bornhove C., Neupert W., Reichert A.S.;
CC "Processing of Mgm1 by the rhomboid-type protease Pcp1 is required for
CC maintenance of mitochondrial morphology and of mitochondrial DNA.";
CC J. Biol. Chem. 278:27781-27788(2003).
CC [7]
CC FUNCTION, SUBCELLULAR LOCATION, VARIANT MD17 LYS-294, MUTAGENESIS OF
CC SER-224, AND IDENTIFICATION OF THE TRANSLATION INITIATION CODON.
CC PubMed=10037792; DOI=10.1083/jcb.144.4.711;
CC Shepard K.A., Yaffe M.P.;
CC "The yeast dynamin-like protein, Mgm1p, functions on the mitochondrial
CC outer membrane to mediate mitochondrial inheritance.";
CC J. Cell Biol. 144:711-720(1999).
CC [8]
CC FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND MUTAGENESIS OF LYS-223;
CC SER-224; THR-244; ARG-824 AND LYS-854
CC PubMed=12566426; DOI=10.1083/jcb.200209015;
CC Wong E.D., Wagner J.A., Scott S.V., Okreglak V., Holewinski T.J.,
CC Cassidy-Stone A., Numata J.;
CC "The intermitochondrial dynamin-related GTPase, Mgm1p, is a component
CC of a protein complex that mediates mitochondrial fusion.";
CC J. Cell Biol. 160:303-311(2003).
CC [9]
CC INTERACTIONS WITH FZO1.
CC PubMed=12808034; DOI=10.1091/mbc.E02-12-0788;
CC Sesaki H., Souhard S.M., Yaffe M.P., Jensen R.E.;
CC "Mgm1p, a dynamin-related GTPase, is essential for fusion of the
CC mitochondrial outer membrane.";
CC Mol. Biol. Cell 14:2342-2356(2003).
CC -1- FUNCTION: Involved in the maintenance of mitochondrial morphology
CC and plays a role in the inheritance of mitochondria. During
CC mitochondrial fusion, Mgm1 coordinates interaction between the
CC inner and outer membranes to promote the formation of the double
CC membrane. Both isoform 1 and isoform 2 are required for full
CC activity. May also function in the transport of mitochondria along
CC extended cytoskeletal tracks.
CC -1- SUBUNIT: Associates with the mitochondria outer membrane proteins
CC FZO1 and UGO1 during mitochondrial fusion.
CC -1- SUBCELLULAR LOCATION: Mitochondrial intermembrane space.
CC -1- PTM: Processed in two steps by mitochondrial proteases. The N-
CC terminal hydrophobic segment is probably cleaved from the
CC precursor by mitochondrial processing peptidase to generate
CC isoform 1. Cleavage of isoform 1 by mitochondrial protease PCP1
CC subsequently yields isoform 2.
CC -1- MISCELLANEOUS: Deletion of Mgm1 causes the mitochondria to
CC fragment and aggregate, and subsequently to lose their
CC mitochondrial DNA and become respiration deficient. The Mdm17
CC mutation results in temperature dependent loss of function.
CC -1- SIMILARITY: Belongs to the dynamin family.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62834; CAA44637.1; ALT_FRAME.
CC EMBL: L07419; -; NOT_ANNOTATED_CDS.
CC EMBL: 275119; CAA99426.1; ALT_INIT.
CC EMBL: 275120; CAA99428.1; ALT_INIT.
CC EMBL: X92441; CAA63174.1; ALT_INIT.
CC EMBL: U11274; AAB59316.1; -.
CC PIR: S33918; S33918.
CC Germonline: 143799; -.
CC DR SGD: S00005737; MGM1.
CC DR GO: GO:0005758; C:mitochondrial intermembrane space; IDA.
CC DR GO: GO:0008053; P:mitochondrial fusion; IDA.
CC DR GO: GO:0000002; P:mitochondrial genome maintenance; IDA.
CC DR GO: GO:0007006; P:mitochondrial membrane organization and bio. .; IDA.
CC DR GO: GO:0000001; P:mitochondrial inheritance; IDA.
CC DR InterPro: IPR001401; Dynamin.
CC DR Pfam: PF00350; Dynamin N.1.
CC DR PRINTS: PR00195; DYNAMIN.
CC DR SMART: SM00053; DYNC.1.
CC DR PROSITE: PS00410; DYNAMIN; 1.
CC KM Direct protein sequencing; GTP-binding; Hydrolase; Mitochondrion;
CC KW Motor protein; Transit peptide.
CC FT TRANSIT 1 60 Mitochondrion.
CC FT CHAIN 61 881 Mgm1 protein, isoform 1.
CC FT DOMAIN 140 881 Mgm1 protein, isoform 2.
CC FT NP_BIND 149 178 Asp-rich (acidic).
CC FT NP_BIND 217 224 GTP (Potential).
CC FT NP_BIND 317 321 GTP (Potential).
CC FT NP_BIND 385 388 GTP (Potential).
CC FT VARIANT 294 294 E -> K (in mdm17).
CC FT MUTAGEN 1 1 M->A: Abolishes translation.
CC FT MUTAGEN 39 39 M->A: No effect on translation.
CC FT MUTAGEN 81 81 M->A: No effect on translation.
CC FT MUTAGEN 92 92 M->A: No effect on translation.
CC FT MUTAGEN 223 223 K->A: Loss of stability.
CC FT MUTAGEN 224 224 S->A: Loss of GTPase activity.
CC FT MUTAGEN 224 224 S->N: Loss of GTPase activity.
CC FT MUTAGEN 244 244 T->A: Loss of GTPase activity.
CC FT MUTAGEN 824 824 R->A: Loss of GED function.
CC FT MUTAGEN 854 854 K->A: Loss of GED function.
CC FT CONFLICT 4 4 S -> T (in Ref. 1).
CC FT CONFLICT 129 129 G -> C (in Ref. 1).
CC FT CONFLICT 170 170 E -> A (in Ref. 2).
CC SQ SEQUENCE 881 AA; 99177 MW; 31524840C5350096 CRC64;
CC -----
CC Query Match 43.9%; Score 47; DB 1; Length 881;
CC Best Local Similarity 72.7%; Pred. No. 1.2e+02;
CC Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC QY 7 KSYFSDQVST 17
CC | : : : : :
CC Db 475 KKYFTNCQVST 485

Search completed: October 27, 2005, 18:55:18
Job time : 67.0465 secs

This Page Blank (100%)

This Page Blank (100%)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:31:33 ; Search time 79.5907 Seconds
(without alignments)
111.765 Million cell updates/sec

Title: US-10-612-090-17
Perfect score: 125
Sequence: 1 SVPNRHHTGVDSLGNFSPILARRV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	23	8	ADJ66744 Human Muc
2	125	100.0	108	8	ADJ66729 Human Muc
3	125	100.0	108	8	ADJ66740 Human Muc
4	125	100.0	174	7	ADM30756 Human can
5	125	100.0	178	5	ABP30979 Predicted
6	125	100.0	178	5	ADA08642 Human O77
7	125	100.0	178	7	ADF08985 Secreted
8	125	100.0	233	5	ABP30978 Extracellular
9	125	100.0	233	7	ADA08641 Human O77
10	125	100.0	233	7	ADF08984 Secreted
11	125	100.0	284	6	ABU54859 Human CA1
12	125	100.0	318	5	ABP31025 Antino aci
13	125	100.0	318	7	ADA08747 Human O77
14	125	100.0	318	7	ADF09090 Secreted
15	125	100.0	367	4	AAE12631 Human gen
16	125	100.0	438	3	AAE12555 Human ova
17	125	100.0	438	4	AAE12623 Human gen
18	125	100.0	438	5	ABP30973 Partial p
19	125	100.0	438	5	ABP30899 0772P c10
20	125	100.0	438	7	ADA08636 Human ova
21	125	100.0	438	7	ADA08543 Human ova
22	125	100.0	438	7	ADF08886 Secreted
23	125	100.0	438	7	ADF08979 Secreted
24	125	100.0	438	7	ADG46173 Human ova
25	125	100.0	439	6	ABU54709 Human CA1

26	125	100.0	526	4	AAE12634 Human gen
27	125	100.0	748	8	ADS94302 CA 125/O7
28	125	100.0	809	8	ADS94303 CA 125/O7
29	125	100.0	833	3	AAE12554 Human ova
30	125	100.0	833	5	ABP30898 Human ova
31	125	100.0	833	7	ADA08542 Human ova
32	125	100.0	833	7	ADF08885 Secreted
33	125	100.0	833	7	ADG46172 Human ova
34	125	100.0	914	4	AAE12552 Human ova
35	125	100.0	914	4	AAE12552 Human ova
36	125	100.0	914	5	ABP30968 Hypocholel
37	125	100.0	914	5	ABP30896 p10
38	125	100.0	914	7	ADA08631 Human O77
39	125	100.0	914	7	ADA08465 Human ova
40	125	100.0	914	7	ADF08974 Secreted
41	125	100.0	914	7	ADF08880 Secreted
42	125	100.0	914	7	ADG46095 Human ova
43	125	100.0	914	8	ADN40451 Human bre
44	125	100.0	1148	4	AAE12552 Human ova
45	125	100.0	1148	4	ABE50283 HOST-1 ov

ALIGNMENTS

RESULT 1
ADJ66744
ID ADJ66744 standard; peptide; 23 AA.
XX
AC ADJ66744;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human Muc1/Muc16 protein-related peptide SegID17.
XX
KW monoclonal antibody; epitope; non-shed extracellular portion;
KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
KW breast cancer; ovarian cancer.
XX
OS Homo sapiens.
XX
PN WO2004005470-A2.
XX
PD 15-JAN-2004.
XX
PF 03-JUL-2003; 2003WO-US020907.
XX
PR 03-JUL-2002; 2002US-0393094P.
XX
PA (IMMU-) IMMUNOGEN INC.
XX
PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
PI Water CA;
XX
DR WPI; 2004-091350/09.
XX
PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
PT treating or monitoring malignancies, such as breast or ovarian cancer.
XX
PS Claim 14; SEQ ID NO 17; 113pp; English.
XX
CC This invention relates to a novel isolated monoclonal antibody that
CC specifically binds to an epitope of a non-shed extracellular portion of a
CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
CC useful for the development of compounds with a cytostatic or for gene
CC therapy. The composition and methods are useful in diagnosing, treating
CC or monitoring malignancies, such as breast or ovarian cancer. The present
CC sequence is that of a peptide which was used for raising antibodies to
CC the extracellular, non-shed region of Muc16 and which is claimed in the
CC specification.
XX
SQ Sequence 23 AA;

Query Match 100.0%; Score 125; DB 8; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.9e-14;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLGNFSPPLARRV 23
 |||||
 DB 1 SVPNRHHTGVDSLGNFSPPLARRV 23

RESULT 2
 ID ADJ66729 standard; protein; 108 AA.
 XX ADJ66729;
 AC ADJ66729;
 XX 06-MAY-2004 (first entry)
 DT
 XX Human Muc16 epitope amino acid sequence.
 DE
 XX monoclonal antibody; epitope; non-shed extracellular portion;
 KM shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KM breast cancer; ovarian cancer.
 XX Homo sapiens.
 OS
 XX WO2004005470-A2.
 PN
 XX 15-JAN-2004.
 PD
 XX 03-JUL-2003; 2003WO-US020907.
 PF
 XX 03-JUL-2002; 2002US-0393094P.
 PR
 XX (IMMU-) IMMUNOGEN INC.
 PA
 XX Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K,
 PI Vater CA;
 PT WPI; 2004-091350/09.
 DR
 XX New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
 PT treating or monitoring malignancies, such as breast or ovarian cancer.
 XX
 PS Claim 13; SEQ ID NO 2; 113pp; English.
 XX
 CC This invention relates to a novel isolated monoclonal antibody that
 CC specifically binds to an epitope of a non-shed extracellular portion of a
 CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
 CC useful for the development of compounds with a cytostatic or for gene
 CC therapy. The composition and methods are useful in diagnosing, treating
 CC or monitoring malignancies, such as breast or ovarian cancer. The present
 CC sequence is that of a Muc epitope of the invention.
 XX
 SQ Sequence 108 AA;

Query Match 100.0%; Score 125; DB 8; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLGNFSPPLARRV 23
 |||||
 DB 37 SVPNRHHTGVDSLGNFSPPLARRV 59

RESULT 3
 ID ADJ66740 standard; protein; 108 AA.
 XX ADJ66740;
 AC ADJ66740;
 XX 06-MAY-2004 (first entry)
 DT
 XX Human Muc16 GST fusion protein amino acid sequence SeqID13.
 DE

XX monoclonal antibody; epitope; non-shed extracellular portion;
 KM shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KM breast cancer; ovarian cancer.
 XX Homo sapiens.
 OS
 XX WO2004005470-A2.
 PN
 XX 15-JAN-2004.
 PD
 XX 03-JUL-2003; 2003WO-US020907.
 PF
 XX 03-JUL-2002; 2002US-0393094P.
 PR
 XX (IMMU-) IMMUNOGEN INC.
 PA
 XX Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K,
 PI Vater CA;
 PT WPI; 2004-091350/09.
 DR
 XX New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
 PT treating or monitoring malignancies, such as breast or ovarian cancer.
 XX
 PS Disclosure; SEQ ID NO 13; 113pp; English.
 XX
 CC This invention relates to a novel isolated monoclonal antibody that
 CC specifically binds to an epitope of a non-shed extracellular portion of a
 CC shed antigen or of human Muc1 or Muc16 protein. The invention may be
 CC useful for the development of compounds with a cytostatic or for gene
 CC therapy. The composition and methods are useful in diagnosing, treating
 CC or monitoring malignancies, such as breast or ovarian cancer. The present
 CC sequence is that of the human Muc16 GST fusion protein which was used in
 CC the exemplification of the invention.
 XX
 SQ Sequence 108 AA;

Query Match 100.0%; Score 125; DB 8; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLGNFSPPLARRV 23
 |||||
 DB 37 SVPNRHHTGVDSLGNFSPPLARRV 59

RESULT 4
 ID ADM30756 standard; protein; 174 AA.
 XX ADM30756;
 AC ADM30756;
 XX 20-MAY-2004 (first entry)
 DT
 XX Human cancer linked protein referred to as 192292 SeqID 4.
 DE
 XX human; cancer-related gene; screening assay; immuno-conjugate;
 KM cytostatic; immunostimulant; cancer; ovarian cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO2003075854-A2.
 PD
 XX 18-SEP-2003.
 PF
 XX 07-MAR-2003; 2003WO-US007147.
 PR
 XX 07-MAR-2002; 2002US-0362527P.
 PA
 XX (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Ebner R, Strovel JW;

XX WPI; 2003-748320/70.
DR N-PSDB; ADM30755.
XX
PT Identifying an agent that modulates the activity of any of three specific
PT cancer-related genes, potentially useful in treating (ovarian) cancer,
PT comprises detecting a difference in expression of the gene in the
PT presence of the agent.
XX
PS Claim 11; SEQ ID NO 4; 57bp; English.
XX
CC This invention relates to a novel method of identifying an agent that
CC modulates the activity of a cancer-related gene. Specifically, it refers
CC to a screening assay to identify potential antitumor agents, as well as
CC methods to assess the cancerous state of a cell. The present invention
CC describes antibodies against the expression products of these cancer-
CC related genes that are capable of targeting a cancerous cell in vivo.
CC Furthermore it provides immuno-conjugates containing such antibodies,
CC which can be used to deliver target therapeutics thereto. Accordingly,
CC these compositions that exhibit cytostatic and immunostimulant activities
CC are useful for preventing or treating cancer in humans and in particular
CC the treatment of ovarian cancer. This polypeptide sequence is a protein
CC encoded by a human cancer-linked gene of the invention.
XX
SQ Sequence 174 AA;
XX
Query Match 100.0%; Score 125; DB 7; Length 174;
Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SVPNRHHTGVDSLGNFSPPLARRV 23
Db 47 SVPNRHHTGVDSLGNFSPPLARRV 69
XX
RESULT 5
ABP30979
ID ABP30979 standard; protein; 178 AA.
XX
AC ABP30979;
XX
DT 02-JUL-2002 (first entry)
XX
DE Predicted extracellular domain of 0772P.
XX
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PI 17-JUL-2000; 2000US-00617747.
XX
PR 10-ANG-2000; 2000US-00635801.
XX
PR 20-SEP-2000; 2000US-00667857.
XX
PR 04-APR-2001; 2001US-00827271.
XX
PR 18-JUN-2001; 2001US-00884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2002-164781/21.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
PS Example 13; Page 361-362; 408bp; English.

XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX
SQ Sequence 178 AA;
XX
Query Match 100.0%; Score 125; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 3.1e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SVPNRHHTGVDSLGNFSPPLARRV 23
Db 106 SVPNRHHTGVDSLGNFSPPLARRV 128
XX
RESULT 6
ADA08642
ID ADA08642 standard; protein; 178 AA.
XX
AC ADA08642;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human 0772P partial protein #12.
XX
XX human; gene therapy; ovarian cancer; cancer.
XX
OS Homo sapiens.
XX
PN US2003091580-A1.
XX
PD 15-MAY-2003.
XX
PF 17-JUL-2001; 2001US-00907969.
XX
PR 18-JUN-2001; 2001US-00884441.
XX
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGN/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
PT New isolated 0772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT cancer.
XX
PS Disclosure; Page 28; 371bp; English.
XX
CC The invention relates to an isolated 0772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma

CC antigen.
 XX Sequence 178 AA;
 SQ

Query Match 100.0%; Score 125; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.1e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLGNFSPFLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLGNFSPFLARRV 128

RESULT 7
 ADF08985 standard; protein; 178 AA.
 XX ADF08985;
 AC
 XX ADF08985;
 AC
 XX 12-FEB-2004 (first entry)
 DT
 XX
 DE Secreted ovarian carcinoma antigen seqid 489.
 XX
 KW gene therapy; protein therapy; vaccine; antibody inhibition;
 KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
 KW secreted ovarian carcinoma antigen.
 XX
 OS Homo sapiens.
 XX
 PN US2003124140-A1.
 PD
 XX 03-JUL-2003.
 PF
 XX 17-JUL-2002; 2002US-00198053.
 PR
 XX 17-DEC-1998; 98US-00215681.
 PR 17-DEC-1998; 98US-00216003.
 PR 23-JUN-1999; 99US-00338933.
 PR 24-SEP-1999; 99US-00404879.
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 PR 17-JUL-2001; 2001US-00907969.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Bangur CS, Retter MW, Fanger GR, Hill P;
 DR WPI; 2003-897152/82.
 XX
 PT Oncogenic nucleic acids useful for the prevention, diagnosis and
 PT treatment of breast cancer.
 XX
 XX Example 13; SEQ ID NO 489; 399pp; English.

The invention describes nucleic acids (I) and the polypeptides (II) they encode. The nucleic acids (I) may be used for preventing, diagnosing and treating diseases related to their aberrant expression i.e. breast cancer. For example, (I) and (II) may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (II) by expressing inactive proteins or to supplement the patients own production of (II). Additionally, (I) may be used to produce (II), by inserting (I) into a host cell and culturing the cell to express the protein (II). (I) and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The host cell may also be used as antigens in the production of antibodies against (II) and in assays to identify modulators of (II)'s expression and activity. The anti-(II) antibodies, agonists and antagonists may be used to regulate expression and activity and as

CC diagnostic agents for detecting the presence of (II) in samples (e.g. by CC immunoassay). This sequence represents a secreted ovarian carcinoma antigen.
 CC
 XX Sequence 178 AA;
 SQ

Query Match 100.0%; Score 125; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.1e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLGNFSPFLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLGNFSPFLARRV 128

RESULT 8
 ABP30978 standard; protein; 233 AA.
 ID ABP30978
 XX ABP30978;
 AC
 XX ABP30978;
 AC
 XX 02-JUL-2002 (first entry)
 DT
 XX
 DE Extracellular and transmembrane regions of 0772P.
 XX
 KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200206317-A2.
 PD
 XX 24-JAN-2002.
 PF
 XX 17-JUL-2001; 2001WO-US022635.
 PR
 XX 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 XX
 DR WPI; 2002-164781/21.
 XX
 PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 XX Example 13; Page 361; 408pp; English.

This invention relates to polypeptides comprising an immunogenic portion of an ovarian carcinoma protein which acts as an immunostimulant and is cytostatic. The polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations and antigen presenting cells that express the polypeptides are useful for stimulating an immune response in a patient and treating ovarian cancer. This sequence represents protein related to the invention

QY 1 SVPNRHHTGVDSLGNFSPFLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLGNFSPFLARRV 128

Query Match 100.0%; Score 125; DB 5; Length 233;
 Best Local Similarity 100.0%; Pred. No. 4.2e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 9
ADA08641
ID ADA08641 standard; protein; 233 AA.
XX
AC ADA08641;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human O772P partial protein #11.
XX
KW human; gene therapy; ovarian cancer; cancer.
XX
OS Homo sapiens.
XX
PN US003091580-A1.
XX
PD 15-MAY-2003.
XX
PF 17-JUL-2001; 2001US-00907969.
XX
PR 18-JUN-2001; 2001US-00884441.
XX
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2003-532352/50.
XX
PT New isolated O772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT cancer.
XX
PS Example 13; Page 28; 371pp; English.
XX
CC The invention relates to an isolated O772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen.
XX
SQ Sequence 233 AA;
XX
QY Query Match 100.0%; Score 125; DB 7; Length 233;
QY Best Local Similarity 100.0%; Pred. No. 4.2e-13;
QY Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVPNRHHGTGVDSLCNFSPPLARRV 23
QY |||||||||||||||||||
QY Db 106 SVPNRHHGTGVDSLCNFSPPLARRV 128
XX
RESULT 10
ADF08984
ID ADF08984 standard; peptide; 233 AA.
XX
AC ADF08984;
XX
DT 12-FEB-2004 (first entry)

```

```

XX
DE Secreted ovarian carcinoma antigen segid 488.
XX
KW gene therapy; protein therapy; vaccine; antibody inhibition;
KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW secreted ovarian carcinoma antigen.
XX
OS Homo sapiens.
XX
PN US003124140-A1.
XX
PD 03-JUL-2003.
XX
PF 17-JUL-2002; 2002US-00198053.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2003-897152/82.
XX
PT Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
PS Example 13; SEQ ID NO 488; 399pp; English.
XX
CC The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patient's own production of (II).
CC Additionally, (I) may be used to produce (II), by inserting (I) into a
CC host cell and culturing the cell to express the protein (II). (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The host cell may also be used as antigens in the production of
CC antibodies against (II) and in assays to identify modulators of (II)'s
CC expression and activity. The anti-(II) antibodies, agonists and
CC antagonists may be used to regulate expression and activity and as
CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC immunoassay). This sequence represents a secreted ovarian carcinoma
CC antigen.
XX
SQ Sequence 233 AA;
XX
QY Query Match 100.0%; Score 125; DB 7; Length 233;
QY Best Local Similarity 100.0%; Pred. No. 4.2e-13;
QY Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVPNRHHGTGVDSLCNFSPPLARRV 23
QY |||||||||||||||||||
QY Db 106 SVPNRHHGTGVDSLCNFSPPLARRV 128
XX
RESULT 11
ABU54859
ID ABU54859 standard; protein; 284 AA.
XX
AC ABU54859;

```

```

XX 12-MAR-2003 (first entry)
XX
XX Human CA125 carboxy terminal domain.
DE
XX Human, CA125; protein repeat; chromosome 19q 13.2; amino terminal domain;
XX amino terminal extension; carboxy terminal domain; vaccine; cancer;
XX ovarian cancer; carcinoma.
XX
XX Homo sapiens.
XX
XX WO200283866-A2.
XX
XX 24-OCT-2002.
XX
XX 12-APR-2002; 2002WO-US011734.
XX
XX 17-APR-2001; 2001US-0284175P.
XX 19-JUN-2001; 2001US-0299380P.
XX 27-SEP-2001; 2001US-00965738.
XX 21-DEC-2001; 2001US-0345180P.
XX
XX (UVAR-) UNIV ARKANSAS.
XX
XX O'Brien T, Beard J, Underwood L;
XX
XX WPI; 2003-093013/08.
XX
XX New CA125 molecules, useful as a gold standard for detecting and
XX monitoring the presence of CA125 antigen which can be used for
XX diagnosing, monitoring or treating patients with cancer or for developing
XX vaccine against cancer.
XX
XX Claim 1; Fig 9b; 694pp; English.
XX
XX The invention relates to a CA125 protein comprising: (a) an extracellular
XX amino terminal domain; (b) an amino terminal extension; (c) a multiple
XX repeat domain; and (d) a carboxy terminal domain. The extracellular amino
XX terminal domain comprises 5 genomic exons, the amino terminal extension
XX comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and
XX the carboxy terminal domain comprises a transmembrane anchor with a short
XX cytoplasmic domain, and further comprises 9 genomic exons. The gene for
XX CA125 is located on human chromosome 19q 13.2. Also included are isolated
XX CA125 repeat domains, nucleic acids (including variants, homologues and
XX degenerate versions) encoding CA125 proteins or repeat units, a vector
XX comprising the nucleic acid, a cultured cell comprising the vector, a
XX method of expressing CA125 antigen in a cell, the amino acid sequences of
XX the CA125 repeat units (or their variants, fragments or sequences 50%
XX identical to them), a purified antibody that selectively binds to an
XX epitope in the receptor-binding domain of CA125 protein, a diagnostic for
XX detecting and monitoring the presence of CA125 antigen (comprising
XX recombinant CA125 having at least one repeat unit of the CA125 repeat
XX domain including epitope binding sites), a therapeutic vaccine to treat
XX mammals with elevated CA125 antigen levels or at risk of developing a
XX disease or disease recurrence associated with elevated CA125 antigen
XX levels (comprising recombinant CA125 repeat domains including epitope
XX binding sites) and an antisense oligonucleotide that inhibits the
XX expression of CA125. The CA125 molecule, particularly the multiple repeat
XX domain are useful as a gold standard for detecting and monitoring the
XX presence of CA125 antigen, which can be used for diagnosing, monitoring
XX or treating patients with ovarian cancer and other carcinomas where CA125
XX is expressed. The molecules are also useful for developing a vaccine
XX against cancer. The present sequence is a CA125 repeat protein
XX
SQ Sequence 284 AA;

```

```

RESULT 12
ABP31025
ID ABP31025 standard; protein, 318 AA.
XX
XX AC ABP31025;
XX
XX 02-JUL-2002 (first entry)
XX
XX Amino acid sequence of the 3' constant region of 0772P.
XX
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
XX
XX Homo sapiens.
XX
XX WO200206317-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001WO-US022635.
XX
XX 17-JUL-2000; 2000US-00617747.
XX 10-AUG-2000; 2000US-00636801.
XX 20-SEP-2000; 2000US-00667857.
XX 04-APR-2001; 2001US-00827271.
XX 18-JUN-2001; 2001US-00884441.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2002-164781/21.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
XX protein or its variants, useful for stimulating an immune response in a
XX patient and treating ovarian cancer.
XX
XX Claim 3; Page 399-400; 408pp; English.
XX
XX This invention relates to polypeptides comprising an immunogenic portion
XX of an ovarian carcinoma protein which acts as an immunostimulant and is
XX cytostatic. The polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations and antigen presenting cells that express
XX the polypeptides are useful for stimulating an immune response in a
XX patient and treating ovarian cancer. This sequence represents protein
XX related to the invention
XX
SQ Sequence 318 AA;

```

```

Query Match 100.0%; Score 125; DB 6; Length 284;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Query Match 100.0%; Score 125; DB 5; Length 318;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
XX US2003091580-A1.
XX
XX 15-MAY-2003.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (MITC/) MITCHAM J L.
XX (KING/) KING G E.
XX (ALGA/) ALGATE P A.
XX (FLIN/) FLING S P.
XX (RETT/) RETTER M W.
XX (FANG/) FANGER G R.
XX (REED/) REED S G.
XX (VEDV/) VEDVICK T S.
XX (CART/) CARTER D.
XX (HILL/) HILL P.
XX (ALBO/) ALBONE E.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
XX New isolated 0772P polypeptides and polynucleotides, useful in gene
XX therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX cancer.
XX
XX Claim 3; SEQ ID NO 594; 371pp; English.
XX
XX The invention relates to an isolated 0772P polypeptide, which has the
XX structure fully defined in the specification. The composition containing
XX the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX or antigen presenting cells are useful for stimulating an immune response
XX and treating ovarian cancer. Detecting the presence of the
XX polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX carcinoma cDNAs and protein cDNAs were identified using microarray
XX technology. The present sequence represents a human ovarian carcinoma
XX antigen.
XX
XX SQ Sequence 318 AA;
XX
XX Query Match 100.0%; Score 125; DB 7; Length 318;
XX Best Local Similarity 100.0%; Pred. No. 6e-13;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SVPNRHHTGVDSLGNFSPILARRV 23
XX |||||||||||||||||||
XX 191 SVPNRHHTGVDSLGNFSPILARRV 213
XX
XX RESULT 14
XX ADF09090
XX ID ADF09090 standard; protein; 318 AA.
XX
XX ADF09090;
XX
XX 12-FEB-2004 (first entry)
XX
XX Secreted ovarian carcinoma antigen seqid 594.
XX
XX gene therapy; protein therapy; vaccine; antibody inhibition;
XX breast cancer; restorative therapy; diagnostic agent; immunoassay;
XX secreted ovarian carcinoma antigen.
XX
XX Homo sapiens.
XX
XX US2003124140-A1.
XX
XX 03-JUL-2003.
XX
```

```
PF 17-JUL-2002; 2002US-00198053.
XX
XX 17-DEC-1998; 98US-00215681.
XX
XX 17-DEC-1998; 98US-00216003.
XX
XX 23-JUN-1999; 99US-00338933.
XX
XX 24-SEP-1999; 99US-00404879.
XX
XX 17-JUL-2000; 2000US-00617747.
XX
XX 10-AUG-2000; 2000US-00636801.
XX
XX 20-SEP-2000; 2000US-00667857.
XX
XX 04-APR-2001; 2001US-00827271.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX (CORI-) CORIXA CORP.
XX
XX
XX Bangur CS, Retter MW, Fanger GR, Hill P;
XX
XX WPI; 2003-897152/82.
XX
XX N-PSDB; ADF09064.
XX
XX
XX Oncogenic nucleic acids useful for the prevention, diagnosis and
XX treatment of breast cancer.
XX
XX Example 16; SEQ ID NO 594; 399pp; English.
XX
XX
XX The invention describes nucleic acids (I) and the polypeptides (II) they
XX encode. The nucleic acids (I) may be used for preventing, diagnosing and
XX treating diseases related to their aberrant expression i.e. breast
XX cancers. For example, (I) and (II) may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of (II) by expressing
XX inactive proteins or to supplement the patient's own production of (II).
XX Additionally, (I) may be used to produce (II), by inserting (I) into a
XX host cell and culturing the cell to express the protein (II). (I) and its
XX complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar nucleic acids in
XX samples, and therefore which patients may be in need of restorative
XX therapy. The host cell may also be used as antigen in the production of
XX antibodies against (II) and in assays to identify modulators of (II)'s
XX expression and activity. The anti-(II) antibodies, agonists and
XX antagonists may be used to regulate expression and activity and as
XX CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
XX CC immunoassay). This sequence represents a secreted ovarian carcinoma
XX antigen.
XX
XX SQ Sequence 318 AA;
XX
XX Query Match 100.0%; Score 125; DB 7; Length 318;
XX Best Local Similarity 100.0%; Pred. No. 6e-13;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SVPNRHHTGVDSLGNFSPILARRV 23
XX |||||||||||||||||||
XX 191 SVPNRHHTGVDSLGNFSPILARRV 213
XX
XX RESULT 15
XX AAE12631
XX ID AAE12631 standard; protein; 367 AA.
XX
XX AAE12631;
XX
XX 03-JAN-2002 (first entry)
XX
XX Human gene 1 encoded secreted protein fragment, SEQ ID NO: 27.
XX
XX Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;
XX respiratory system disorder; asthma; haematopoietic disorder; skin aging;
XX immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
XX rheumatoid arthritis; inflammation; neurological disorder; septic shock;
XX Alzheimer's disease; Parkinson's disease; diabetes; angiodenic disorder;
XX atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
XX epithelial cell proliferation; transplantation; chemocaxis; infection;
XX
```

KM food additive; wound healing; endocrine disorder; kidney disorder;
 KW gene therapy; cytostatic.
 OS Homo sapiens.
 PN WO200170804-A1.
 PD 27-SEP-2001.
 PF 16-MAR-2001; 2001WO-US008585.
 PR 17-MAR-2000; 2000US--0190076P.
 PR 23-AUG-2000; 2000US-0227009P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Birse CE, Rosen CA;
 DR WPI; 2001-639119/73.
 XX Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
 PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
 PT metastases.
 PS Disclosure; Page 9; 427pp; English.
 CC AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
 CC ovarian cancer associated protein (collectively known as ovarian cancer
 CC antigens) genes, and AAR12623-AAR12629 represent the proteins they
 CC encode. AAR12630-AAR12638 represent human ovarian cancer antigenic
 CC fragments or variants. Ovarian cancer antigens and their corresponding
 CC DNAs are used in the prevention, diagnosis and treatment of diseases
 CC associated with their inappropriate expression. These disorders include
 CC proliferative disorders, cancer, tumours, respiratory system disorders,
 CC asthma, haematopoietic disorders, diseases of the immune system, AIDS,
 CC skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
 CC inflammation, allergies, neurological disorders (e.g., Alzheimer's
 CC disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, endocrine disorders and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, to identify their cognate ligands or binding
 CC partners, in chemotaxis and can be used as a food additive. Antibodies
 CC specific for a protein of the invention can be used in alleviating
 CC symptoms associated with the disorders mentioned above and in diagnostic
 CC immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
 CC the invention is used in gene therapy. The present sequence represents a
 CC human ovarian cancer antigenic fragment of the invention
 XX
 SQ Sequence 367 AA;
 Query Match 100.0%; Score 125; DB 4; Length 367;
 Best Local Similarity 100.0%; Pred. No. 7,1e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVPNRHTGTVDLNCNFSPLARRV 23
 ||||||||||||||||||
 DB 311 SVPNRHTGTVDLNCNFSPLARRV 333

Search completed: October 27, 2005, 18:44:25
 Job time : 80.5907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2005, 17:03:42 ; Search time 20.0047 Seconds
(without alignments)
85.826 Million cell updates/sec

Title: US-10-612-090-17
Perfect score: 125
Sequence: 1 SVNRRHHTGVDSLNCNFSPLARRV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	125	100.0	438	4	US-09-404-879A-390	Sequence 390, App
2	125	100.0	438	4	US-09-667-857-390	Sequence 390, App
3	125	100.0	833	4	US-09-404-879A-389	Sequence 389, App
4	125	100.0	833	4	US-09-667-857-389	Sequence 389, App
5	125	100.0	914	4	US-09-404-879A-312	Sequence 312, App
6	125	100.0	914	4	US-09-338-933-312	Sequence 312, App
7	125	100.0	914	4	US-09-667-857-312	Sequence 312, App
8	51	40.8	942	3	US-09-171-461-12	Sequence 12, Appl
9	51	40.8	942	4	US-09-970-711-12	Sequence 12, Appl
10	49.5	39.6	113	4	US-09-884-570-9	Sequence 9, Appl
11	49.5	39.6	113	4	US-10-157-457A-12	Sequence 12, Appl
12	49.5	39.6	113	4	US-10-157-457A-12	Sequence 12, Appl
13	46.5	37.2	230	4	US-09-252-991A-18747	Sequence 18747, A
14	46	36.8	808	4	US-09-252-991A-19470	Sequence 19470, A
15	45	36.0	420	4	US-09-527-084A-6	Sequence 6, Appl
16	44	35.2	168	4	US-09-248-796A-14756	Sequence 14756, A
17	44	35.2	212	4	US-09-328-352-6982	Sequence 6982, Ap
18	43.5	34.8	211	4	US-09-352-991A-20225	Sequence 20225, A
19	43	34.4	80	4	US-09-248-796A-21674	Sequence 21674, A
20	43	34.4	246	4	US-09-252-991A-32673	Sequence 32673, A
21	43	34.4	410	4	US-09-552-991A-26654	Sequence 26654, A
22	43	34.4	734	4	US-10-237-551-26	Sequence 26, Appl
23	43	34.0	22	3	US-08-318-794-30	Sequence 30, Appl
24	42.5	34.0	22	3	US-08-470-106-30	Sequence 30, Appl
25	42.5	34.0	427	4	US-09-489-039A-10737	Sequence 10737, A
26	42.5	34.0	427	4	US-09-489-039A-10737	Sequence 10737, A
27	42	33.6	419	4	US-09-489-039A-13188	Sequence 13188, A

28	42	33.6	843	4	US-09-489-039A-7545	Sequence 7545, Ap
29	41.5	33.2	391	4	US-09-328-352-4861	Sequence 4861, Ap
30	41.5	33.2	1025	4	US-09-834-309-5	Sequence 5, Appl
31	41.5	33.2	1151	4	US-09-177-165A-31	Sequence 31, Appl
32	41	32.8	320	4	US-10-014-268-30	Sequence 30, Appl
33	41	32.8	321	4	US-07-945-288-6	Sequence 6, Appl
34	41	32.8	321	1	US-08-462-831-6	Sequence 6, Appl
35	41	32.8	321	1	US-08-461-809-6	Sequence 6, Appl
36	41	32.8	321	1	US-08-461-441-6	Sequence 6, Appl
37	41	32.8	321	2	US-08-482-142-6	Sequence 6, Appl
38	41	32.8	321	2	US-08-478-572-6	Sequence 6, Appl
39	41	32.8	321	3	US-08-484-296-6	Sequence 6, Appl
40	41	32.8	321	5	PCT-US93-08518-6	Sequence 6, Appl
41	41	32.8	395	4	US-09-252-991A-30497	Sequence 30497, A
42	41	32.8	490	3	US-09-099-041A-26	Sequence 26, Appl
43	41	32.8	490	3	US-09-245-281-26	Sequence 26, Appl
44	41	32.8	490	4	US-09-207-359B-26	Sequence 26, Appl
45	41	32.8	490	4	US-09-340-620A-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-404-879A-390
; Sequence 390, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404, 879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 390
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-390
Query Match 100.0%; Score 125; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.6e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVNRRHHTGVDSLNCNFSPLARRV 23
Db 311 SVNRRHHTGVDSLNCNFSPLARRV 333
RESULT 2
US-09-667-857-390
; Sequence 390, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedyck, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667, 857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 390
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-857-390

```

```

Query Match          100.0%; Score 125; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 6,66-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SVPNRHHTGVSLCNFSPPLARRV 23
      |||
DB      311 SVPNRHHTGVSLCNFSPPLARRV 333

```

```

RESULT 3
US-09-404-879A-389
; Sequence 389, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-389

```

```

Query Match          100.0%; Score 125; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 1,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SVPNRHHTGVSLCNFSPPLARRV 23
      |||
DB      706 SVPNRHHTGVSLCNFSPPLARRV 728

```

```

RESULT 4
US-09-667-857-389
; Sequence 389, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-857-389

```

```

Query Match          100.0%; Score 125; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 1,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SVPNRHHTGVSLCNFSPPLARRV 23
      |||
DB      706 SVPNRHHTGVSLCNFSPPLARRV 728

```

```

RESULT 5
US-09-404-879A-312
; Sequence 312, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-404-879A-312

```

```

Query Match          100.0%; Score 125; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1,5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SVPNRHHTGVSLCNFSPPLARRV 23
      |||
DB      787 SVPNRHHTGVSLCNFSPPLARRV 809

```

```

RESULT 6
US-09-338-933-312
; Sequence 312, Application US/09338933
; Patent No. 6468931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-338-933-312

```

```

Query Match          100.0%; Score 125; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1,5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SVPNRHHTGVSLCNFSPPLARRV 23
      |||
DB      787 SVPNRHHTGVSLCNFSPPLARRV 809

```

```

RESULT 7
US-09-667-857-312
; Sequence 312, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:

```

APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon R.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Marc W.
APPLICANT: Renger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C5
CURRENT APPLICATION NUMBER: US/09/667,857
CURRENT FILING DATE: 2000-09-20
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 312
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapien
US-09-667-857-312

Query Match 100.0%; Score 125; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVNRHHTGVDSLNCNFSPLARV 23
DB 787 SVNRHHTGVDSLNCNFSPLARV 809

RESULT 8
US-09-171-461-12
Sequence 12, Application US/09171461
Patent No. 6335016
GENERAL INFORMATION:
APPLICANT: Baker, Adam
APPLICANT: Cotten, Matthew
APPLICANT: Chiocca, Susanna
APPLICANT: Kurzbauer, Robert
APPLICANT: Schaffner, Gotthold
TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
FILE REFERENCE: 0652.1800000
CURRENT APPLICATION NUMBER: US/09/171,461
CURRENT FILING DATE: 1999-01-12
EARLIER APPLICATION NUMBER: PCT/EP97/01944
EARLIER FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 942
TYPE: PRT
ORGANISM: CELO Virus
FEATURE:
OTHER INFORMATION: Position: 18289..21117 /gene: L3 /product: L3 hexon
US-09-171-461-12

Query Match 40.8%; Score 51; DB 3; Length 942;
Best Local Similarity 45.0%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLNCNFSPLAR 22
DB 757 PDRHYFHYDFLRNFDPMRQ 776

RESULT 9
US-09-970-711-12
Sequence 12, Application US/09970711
Patent No. 6773709
GENERAL INFORMATION:
APPLICANT: Baker, Adam
APPLICANT: Cotten, Matthew

APPLICANT: Chiocca, Susanna
APPLICANT: Kurzbauer, Robert
APPLICANT: Schaffner, Gotthold
TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
FILE REFERENCE: 0652.1800001
CURRENT APPLICATION NUMBER: US/09/970,711
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/171,461
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: PCT/EP97/01944
PRIOR FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 942
TYPE: PRT
ORGANISM: CELO Virus
FEATURE:
OTHER INFORMATION: Position: 18289..21117 /gene: L3 /product: L3 hexon
US-09-970-711-12

Query Match 40.8%; Score 51; DB 4; Length 942;
Best Local Similarity 45.0%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLNCNFSPLAR 22
DB 757 PDRHYFHYDFLRNFDPMRQ 776

RESULT 10
US-09-884-570-9
Sequence 9, Application US/09884570
Patent No. 6566333
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
FILE REFERENCE: D6294
CURRENT APPLICATION NUMBER: US/09/884,570
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/212,271
PRIOR FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 9
LENGTH: 113
TYPE: PRT
ORGANISM: unknown
FEATURE:
LOCATION: 38..165
OTHER INFORMATION: amino acid sequence of region of Lect2
OTHER INFORMATION: protein homologous to repeat sequences of N
OTHER INFORMATION: and C terminal ends of the mim-1 protein
US-09-884-570-9

Query Match 39.6%; Score 49.5; DB 4; Length 113;
Best Local Similarity 37.5%; Pred. No. 1.8;
Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 1;

QY 5 RHHTGVDSLNCN-----FSPPLARV 23
DB 13 RHHPGVDSLCDGSGVYVAPFTGKI 36

RESULT 11
US-10-157-457A-9
Sequence 9, Application US/10157457A
Patent No. 6800610
GENERAL INFORMATION:
APPLICANT: Williams, John P.
APPLICANT: McDonald, Jay M.
APPLICANT: McKenna, Margaret A.

```

; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294CIP
; CURRENT APPLICATION NUMBER: US/10/157,457A
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 09/884,570
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 9
; LENGTH: 113
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; LOCATION: 38..165
; OTHER INFORMATION: amino acid sequence of region of Lect2
; OTHER INFORMATION: protein homologous to repeat sequences of N
; OTHER INFORMATION: and C terminal ends of the mim-1 protein
US-10-157-457A-9

Query Match          39.6%; Score 49.5; DB 4; Length 113;
Best Local Similarity 37.5%; Pred. No. 1.8;
Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 1;

Oy      5 RHHTGVDSLGN-----FSPPLARR 23
        ||| ||| ||| : : : : :
Db      13 RHHPGVDLCSGDSVYVAPFTGKI 36

RESULT 12
US-10-157-457A-12
; Sequence 12, Application US/10157457A
; Patent No. 6800610
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294CIP
; CURRENT APPLICATION NUMBER: US/10/157,457A
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 09/884,570
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; OTHER INFORMATION: full length amino acid sequence of mim-1 protein
US-10-157-457A-12

Query Match          39.6%; Score 49.5; DB 4; Length 113;
Best Local Similarity 37.5%; Pred. No. 1.8;
Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 1;

Oy      5 RHHTGVDSLGN-----FSPPLARR 23
        ||| ||| ||| : : : : :
Db      13 RHHPGVDLCSGDSVYVAPFTGKI 36

RESULT 13
US-09-252-991A-18747
; Sequence 18747, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18747
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18747

Query Match          37.2%; Score 46.5; DB 4; Length 230;
Best Local Similarity 45.5%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Oy      1 SVPNRHTGVDSLGNFSPPLARR 22
        | : | | | | : : : : |
Db      50 SAPSRAHTGSSRRCS-APTSSR 70

RESULT 14
US-09-252-991A-19470
; Sequence 19470, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19470
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19470

Query Match          36.8%; Score 46; DB 4; Length 808;
Best Local Similarity 45.0%; Pred. No. 58;
Matches 9; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

Oy      5 RHHTGVDSL--GNFSPPLARR 22
        ||| ||| : : : : |
Db      550 RHHDGKDTFVCDPQGVCR 569

RESULT 15
US-09-527-084A-6
; Sequence 6, Application US/09527084A
; Patent No. 6696560
; GENERAL INFORMATION:
; APPLICANT: Durfee, Tim
; APPLICANT: Feller, Heidi
; APPLICANT: Gruissem, Wilhelm
; APPLICANT: Jenkins, Susan
; APPLICANT: Roe, Judith
; APPLICANT: Zambryewski, Patricia
; TITLE OF INVENTION: Alteration of Plant Meristem Function by Manipulation of the
; TITLE OF INVENTION: Retinoblastoma-Like Plant RB Gene
; FILE REFERENCE: S-92981
; CURRENT APPLICATION NUMBER: US/09/527,084A
; CURRENT FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/125,229
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Zea mays
US-09-527-084A-6

```

Query Match 36.0%; Score 45; DB 4; Length 420;
 Best Local Similarity 45.0%; Pred. No. 41;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 PNRHRTGVDSLNCNFSPLARR 22
 |||:|
 DB 397 PNGGSTLDPAAFSPISKR 416

Search completed: October 27, 2005, 17:34:34
 Job time : 21.0047 secs

THE NEW YORK PUBLIC LIBRARY
ASTOR LENOX TILDEN FOUNDATION
125 WEST 47TH STREET
NEW YORK, N. Y. 10036

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:30:08 ; Search time 73.6 Seconds
(without alignments)
130.638 Million cell updates/sec

Title: US-10-612-090-17
Perfect score: 125
Sequence: 1 SVPNRHTGVDSLGNFSPILARRV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10A_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUBCOMB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	23	US-10-612-090-17	Sequence 17, Appl
2	125	100.0	108	US-10-612-090-2	Sequence 2, Appl
3	125	100.0	108	US-10-612-090-13	Sequence 13, Appl
4	125	100.0	174	US-10-383-368-4	Sequence 4, Appl
5	125	100.0	178	US-09-884-441-489	Sequence 489, App
6	125	100.0	178	US-09-907-969-489	Sequence 489, App
7	125	100.0	178	US-10-198-053-489	Sequence 489, App
8	125	100.0	178	US-10-860-790-489	Sequence 489, App
9	125	100.0	233	US-09-884-441-488	Sequence 488, App
10	125	100.0	233	US-09-907-969-488	Sequence 488, App
11	125	100.0	233	US-10-198-053-488	Sequence 488, App

12	125	100.0	233	17	US-10-860-790-488	Sequence 488, App
13	125	100.0	284	10	US-09-965-738-300	Sequence 300, App
14	125	100.0	318	10	US-09-907-969-594	Sequence 594, App
15	125	100.0	318	14	US-10-198-053-594	Sequence 594, App
16	125	100.0	318	17	US-10-860-790-594	Sequence 594, App
17	125	100.0	367	15	US-10-333-900-27	Sequence 27, Appl
18	125	100.0	396	18	US-10-858-412-225	Sequence 225, App
19	125	100.0	438	9	US-09-884-441-390	Sequence 390, App
20	125	100.0	438	9	US-09-884-441-483	Sequence 483, App
21	125	100.0	438	10	US-09-907-969-390	Sequence 390, App
22	125	100.0	438	10	US-09-907-969-483	Sequence 483, App
23	125	100.0	438	10	US-09-827-271-390	Sequence 390, App
24	125	100.0	438	14	US-10-198-053-390	Sequence 390, App
25	125	100.0	438	14	US-10-198-053-483	Sequence 483, App
26	125	100.0	438	15	US-10-333-900-30	Sequence 30, Appl
27	125	100.0	438	17	US-10-860-790-390	Sequence 390, App
28	125	100.0	438	17	US-10-860-790-483	Sequence 483, App
29	125	100.0	439	10	US-09-965-738-148	Sequence 148, App
30	125	100.0	456	18	US-10-858-412-226	Sequence 226, App
31	125	100.0	526	15	US-10-333-900-30	Sequence 30, Appl
32	125	100.0	545	14	US-10-243-243A-4	Sequence 4, Appl
33	125	100.0	583	14	US-10-142-515-4	Sequence 4, Appl
34	125	100.0	748	17	US-10-687-035-1	Sequence 1, Appl
35	125	100.0	809	17	US-10-687-035-2	Sequence 2, Appl
36	125	100.0	833	9	US-09-884-441-389	Sequence 389, App
37	125	100.0	833	10	US-09-907-969-389	Sequence 389, App
38	125	100.0	833	10	US-09-827-271-389	Sequence 389, App
39	125	100.0	833	14	US-10-198-053-389	Sequence 389, App
40	125	100.0	833	17	US-10-860-790-389	Sequence 389, App
41	125	100.0	914	9	US-09-778-320-206	Sequence 206, App
42	125	100.0	914	9	US-09-910-689-206	Sequence 206, App
43	125	100.0	914	9	US-09-884-441-312	Sequence 312, App
44	125	100.0	914	9	US-09-884-441-478	Sequence 478, App
45	125	100.0	914	10	US-09-907-969-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-10-612-090-17
; Sequence 17, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-17

Query Match 100.0%; Score 125; DB 15; Length 23;
Best Local Similarity 100.0%; Pred. No. 4,3-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHTGVDSLGNFSPILARRV 23
DB 1 SVPNRHTGVDSLGNFSPILARRV 23

RESULT 2
US-10-612-090-2
; Sequence 2, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:

```
/ APPLICANT: Immunogen, Inc.
/ TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
/ FILE REFERENCE: A8340
/ CURRENT APPLICATION NUMBER: US/10/612,090
/ PRIOR FILING DATE: 2003-07-03
/ PRIOR APPLICATION NUMBER: US 60/393,094
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2
/ LENGTH: 108
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-612-090-2

Query Match
Best Local Similarity 100.0%; Score 125; DB 15; Length 108;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SVPNRHTGVDSLGNFSPFLARRV 23
Db 37 SVPNRHTGVDSLGNFSPFLARRV 59

RESULT 3
US-10-612-090-13
/ Sequence 13, Application US/10612090
/ Publication No. US20040057952A1
/ GENERAL INFORMATION:
/ APPLICANT: Immunogen, Inc.
/ TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
/ FILE REFERENCE: A8340
/ CURRENT APPLICATION NUMBER: US/10/612,090
/ PRIOR FILING DATE: 2003-07-03
/ PRIOR APPLICATION NUMBER: US 60/393,094
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 13
/ LENGTH: 108
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Fusion protein
/ NAME/KEY: MISC FEATURE
/ LOCATION: (1)..(1)
/ OTHER INFORMATION: Glutathione S-transferase fusion site
US-10-612-090-13

Query Match
Best Local Similarity 100.0%; Score 125; DB 15; Length 108;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SVPNRHTGVDSLGNFSPFLARRV 23
Db 37 SVPNRHTGVDSLGNFSPFLARRV 59

RESULT 4
US-10-383-368-4
/ Sequence 4, Application US/10383368
/ Publication No. US20040002091A1
/ GENERAL INFORMATION:
/ APPLICANT: Young, Paul E.
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Strovel, Jeffrey W.
/ TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
/ FILE REFERENCE: 689290-123
/ CURRENT APPLICATION NUMBER: US/10/383,368
/ PRIOR FILING DATE: 2003-03-07
/ PRIOR APPLICATION NUMBER: US 60/362,527
/ PRIOR FILING DATE: 2002-03-07
```

```
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 174
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-383-368-4

Query Match
Best Local Similarity 100.0%; Score 125; DB 15; Length 174;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SVPNRHTGVDSLGNFSPFLARRV 23
Db 47 SVPNRHTGVDSLGNFSPFLARRV 69

RESULT 5
US-09-884-441-489
/ Sequence 489, Application US/09884441
/ Patent No. US20020119158A1
/ GENERAL INFORMATION:
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Carter, Darick
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.462C7
/ CURRENT APPLICATION NUMBER: US/09/884,441
/ PRIOR FILING DATE: 2001-06-18
/ NUMBER OF SEQ ID NOS: 489
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 489
/ LENGTH: 178
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-884-441-489

Query Match
Best Local Similarity 100.0%; Score 125; DB 9; Length 178;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SVPNRHTGVDSLGNFSPFLARRV 23
Db 106 SVPNRHTGVDSLGNFSPFLARRV 128

RESULT 6
US-09-907-969-489
/ Sequence 489, Application US/09907969
/ Publication No. US20030091580A1
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Reltter, Marc W.
/ APPLICANT: Fanger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darick
/ APPLICANT: Hill, Paul
/ APPLICANT: Albone, Earl
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.462C8
/ CURRENT APPLICATION NUMBER: US/09/907,969
/ PRIOR FILING DATE: 2001-07-17
/ NUMBER OF SEQ ID NOS: 596
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 489
/ LENGTH: 178
/ TYPE: PRT
/ ORGANISM: Homo sapiens
```


US-09-907-969-489

Query Match 100.0%; Score 125; DB 10; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.7e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLCNFSPLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLCNFSPLARRV 128

RESULT 7

US-10-198-053-489
 ; Sequence 489, Application US/10198053
 ; Publication No. US20030124140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Hill, Paul
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.462C9
 ; CURRENT APPLICATION NUMBER: US/10/198.053
 ; CURRENT FILING DATE: 2002-07-17
 ; NUMBER OF SEQ ID NOS: 624
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 489
 ; LENGTH: 178
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-198-053-489

Query Match 100.0%; Score 125; DB 14; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.7e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLCNFSPLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLCNFSPLARRV 128

RESULT 8

US-10-860-790-489
 ; Sequence 489, Application US/10860790
 ; Publication No. US20050031634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Hill, Paul
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.462C11
 ; CURRENT APPLICATION NUMBER: US/10/860.790
 ; CURRENT FILING DATE: 2004-06-02
 ; NUMBER OF SEQ ID NOS: 624
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 489
 ; LENGTH: 178
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-860-790-489

Query Match 100.0%; Score 125; DB 17; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.7e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLCNFSPLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLCNFSPLARRV 128

RESULT 9

US-09-884-441-488
 ; Sequence 488, Application US/09884441
 ; Patent No. US20020119158A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Carter, Darrick
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.462C7
 ; CURRENT APPLICATION NUMBER: US/09/884.441
 ; CURRENT FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 489
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 488
 ; LENGTH: 233
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-884-441-488

Query Match 100.0%; Score 125; DB 9; Length 233;
 Best Local Similarity 100.0%; Pred. No. 4.9e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLCNFSPLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLCNFSPLARRV 128

RESULT 10

US-09-907-969-488
 ; Sequence 488, Application US/09907969
 ; Publication No. US20030091580A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Fling, Steven P.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Fanger, Gary Richard
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Hill, Paul
 ; APPLICANT: Albone, Earl
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.462C8
 ; CURRENT APPLICATION NUMBER: US/09/907.969
 ; CURRENT FILING DATE: 2001-07-17
 ; NUMBER OF SEQ ID NOS: 596
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 488
 ; LENGTH: 233
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-907-969-488

Query Match 100.0%; Score 125; DB 10; Length 233;
 Best Local Similarity 100.0%; Pred. No. 4.9e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVPNRHHTGVDSLCNFSPLARRV 23
 |||||
 DB 106 SVPNRHHTGVDSLCNFSPLARRV 128

RESULT 11

US-10-198-053-488
 ; Sequence 488, Application US/10198053
 ; Publication No. US20030124140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bangur, Chaitanya S.

```

; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-488

Query Match          100.0%; Score 125; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVPNRHHGTGVDSLCNFSPFLARRV 23
Db 106 SVPNRHHGTGVDSLCNFSPFLARRV 128

RESULT 12
US-10-860-790-488
; Sequence 488, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-790-488

Query Match          100.0%; Score 125; DB 17; Length 233;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVPNRHHGTGVDSLCNFSPFLARRV 23
Db 106 SVPNRHHGTGVDSLCNFSPFLARRV 128

RESULT 13
US-09-965-738-300
; Sequence 300, Application US/09965738
; Publication No. US20030143667A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy
; TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic a
; FILE REFERENCE: 40715-258841
; CURRENT APPLICATION NUMBER: US/09/965,738
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/284,175
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 300
; LENGTH: 284

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-738-300

Query Match          100.0%; Score 125; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVPNRHHGTGVDSLCNFSPFLARRV 23
Db 157 SVPNRHHGTGVDSLCNFSPFLARRV 179

RESULT 14
US-09-907-969-594
; Sequence 594, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-09-907-969-594

Query Match          100.0%; Score 125; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVPNRHHGTGVDSLCNFSPFLARRV 23
Db 191 SVPNRHHGTGVDSLCNFSPFLARRV 213

RESULT 15
US-10-198-053-594
; Sequence 594, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 318

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-10-198-053-594

```

```

Query Match          100.0%; Score 125; DB 14; Length 318;
Best Local Similarity 100.0%; Pred No. 6, 8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  SVPNRHHTGVDSL CNFSP LARRV 23
          |||||
Db      191 SVPNRHHTGVDSL CNFSP LARRV 213

```

Search completed: October 27, 2005, 18:31:47
 Job time : 74.6 secs

Tim Pegg Book (1970)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 18:13:49 ; Search time 14.2279 Seconds
(without alignments)
155.538 Million cell updates/sec

Title: US-10-612-090-17

Perfect score: 125

Sequence: 1 SVPNRHHTGVDSLNCNFSPLARRV 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	41.6	390	2	F84196	hypothetical prote
2	49	39.2	209	2	T44653	capsular polysacch
3	49	39.2	266	2	T15066	hypothetical prote
4	48	38.4	462	2	T25726	hypothetical prote
5	47	37.6	294	2	AF1218	Salmonella typhim
6	47	37.6	294	2	A11571	Regulatory protein
7	46	36.8	774	2	D83507	hypothetical prote
8	45	36.0	285	2	A82609	hypothetical prote
9	45	36.0	366	2	T25178	hypothetical prote
10	45	36.0	420	2	T01172	G1/S transilition co
11	43.5	34.8	332	2	C83933	penicillin-binding
12	43.5	34.8	403	2	H64861	hypothetical prote
13	43	34.4	325	2	B82201	conserved hypotet
14	43	34.4	335	2	G64817	Probable membrane
15	43	34.4	344	1	WMBR31	38K protein - huma
16	43	34.4	530	2	T18596	hypothetical prote
17	43	34.4	735	1	WMBRT5	U15 protein - hum
18	43	34.4	1018	2	T19693	hypothetical prote
19	43	34.4	2160	2	T20241	hypothetical prote
20	42.5	34.0	160	2	I84444	eosinophil-derived
21	42	33.6	117	1	ERAD01	early E3 13k glyco
22	42	33.6	244	1	P2MMBB	conserved hypotet
23	42	33.6	810	1	P2MMBB	2a protein - broad
24	42	33.6	813	2	T21192	hypothetical prote
25	41.5	33.2	122	2	E70738	hypothetical prote
26	41.5	33.2	232	2	B82140	probable C4-dicard
27	41.5	33.2	508	2	S51809	phosphoprotein pho
28	41.5	33.2	591	2	H86267	probable protein p
29	41.5	33.2	982	2	T43699	DNA mismatch repai

30	41.5	33.2	1025	1	A43526	complement C3d/Bps
31	41.5	33.2	1151	2	A41529	GRII protein - yea
32	41.5	33.2	1239	2	T06143	disease resistance
33	41.5	33.2	3944	2	T19997	hypothetical prote
34	41	32.8	152	2	F84642	hypothetical prote
35	41	32.8	168	2	T28745	hypothetical prote
36	41	32.8	280	2	T17116	protein kinase cdc
37	41	32.8	319	2	A61500	allergen Der f I p
38	41	32.8	335	2	G85589	hypothetical prote
39	41	32.8	335	2	E90739	hypothetical prote
40	41	32.8	350	2	E84109	potassium uptake p
41	41	32.8	410	2	S77844	alanine-tRNA ligase
42	41	32.8	453	2	A96688	hypothetical prote
43	41	32.8	487	2	T34868	probable glutamate
44	41	32.8	589	2	AD3263	hypothetical prote
45	41	32.8	711	2	P65201	phosphoenolpyruvat

ALIGNMENTS

RESULT 1

F84196 hypothetical protein Vng0378c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: F84196

R:Ng, W.V.; Kennedy, S.P.; Mahataas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laesky, S.

; Leithauser, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor

Jung, K.H.; Alam, M.; Freilich, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.; Li

A>Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: F84196

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-390 <STO>

A:Cross-references: UNIPROT:Q9H570; GB:AE004437; NID:G10579994; PIDN:AA618938.1; GSPDB:GT

C:Genetics:

A:Gene: VNG0378C

Query Match 41.6%; Score 52; DB 2; Length 390;

Best Local Similarity 56.2%; Pred. No. 1.3;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLNCNFSPL 18

Db 41 PNAHNAVVDLSRSLVAP 56

RESULT 2

T44653 capsular polysaccharide biosynthesis protein cpsE [imported] - Streptococcus agalactiae

C:Species: Streptococcus agalactiae

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C:Accession: T44653

R:Chaffin, D.O.; Yim, H.H.; Beres, S.B.; Sweet, E.S.; Nittayaajarn, A.; Rubens, C.E.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z22821

A:Accession: T44653

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-209 <CHA>

A:Cross-references: UNIPROT:Q9RBP8; UNIPROT:Q8DZ55; EMBL:AF163833; PIDN:AA053076.1

A:Experimental source: strain COH1; serotype III

C:Genetics:

A:Gene: cpsE

Query Match 39.2%; Score 49; DB 2; Length 209;

Best Local Similarity 60.0%; Pred. No. 2;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:34:49 ; Search time 66.8605 Seconds
(without alignments)
176.155 Million cell updates/sec

Title: US-10-612-090-17
Perfect score: 125
Sequence: 1 SVPNRHHTGVDSLGNFSPILARRV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	1148	2 Q9H7S7	Q9H7S7 homo sapien
2	125	100.0	6995	2 Q96RK2	Q96RK2 homo sapien
3	125	100.0	22152	2 Q8WXI7	Q8WXI7 homo sapien
4	98	78.4	258	2 Q9DHI1	Q9DHI1 mus musculi
5	52	41.6	141	2 Q96C55	Q96C55 rhizobium l
6	52	41.6	390	1 Y378 HALN1	Q9H670 halobacteri
7	51	40.8	942	1 HEX ADEG1	P42671 avian adeno
8	51	40.8	942	2 Q96585	Q96585 avian adeno
9	51	40.8	942	2 Q7MSF4	Q7MSF4 avian adeno
10	50.5	40.4	785	2 Q614N2	Q614N2 oryza sativ
11	49.5	39.6	151	1 LC72 MOUSE	Q88803 mus musculi
12	49.5	39.6	151	2 Q8K181	Q8K181 mus musculi
13	49	39.2	207	2 Q9XDT0	Q9XDT0 streptococc
14	49	39.2	209	2 Q9RPB8	Q9RPB8 streptococc
15	49	39.2	209	2 Q8DZES	Q8DZES streptococc
16	49	39.2	209	2 Q8E503	Q8E503 streptococc
17	49	39.2	266	2 Q9T2Z9	Q9T2Z9 caenorhabdi
18	48.5	38.8	202	2 Q9CKQ5	Q9CKQ5 pasteurella
19	48.5	38.8	521	2 Q6FUK9	Q6FUK9 candida gla
20	48.5	38.8	869	2 Q6MBE5	Q6MBE5 paracitampd
21	48	38.4	462	1 HMCS CAEEL	P54871 caenorhabdi
22	47	37.6	198	2 Q9L180	Q9L180 streptomyce
23	47	37.6	294	2 Q8Y7W9	Q8Y7W9 listeria mo
24	47	37.6	294	2 Q92CQ7	Q92CQ7 listeria in
25	47	37.6	294	2 Q720T1	Q720T1 listeria mo
26	47	37.6	541	1 FTCD HUMAN	Q93954 h formimid
27	47	37.6	554	2 Q86VQ3	Q86VQ3 homo sapien
28	46.5	37.2	1750	2 Q9RPF5	Q9RPF5 stigmatella
29	46	36.8	405	2 Q98146	Q98146 rhizobium l
30	46	36.8	440	2 Q8FT90	Q8FT90 corynebacte
31	46	36.8	496	2 Q93J15	Q93J15 streptomyce

32	46	36.8	774	2 Q914M0	Q914M0 pseudomonas
33	46	36.8	873	1 SYA WIGBR	Q8d2w8 wigleswort
34	46	36.8	1981	2 Q9E9R6	Q9E9R6 hop latent
35	45.5	36.4	1045	2 Q46545	Q46545 ovie aries
36	45.5	36.4	1343	1 TIM DROVI	Q17482 drosophila
37	45	36.0	163	2 Q8DLB0	Q8DLB0 synchococc
38	45	36.0	242	2 Q852C9	Q852C9 oryza sativ
39	45	36.0	245	2 Q7N564	Q7N564 photorhabdu
40	45	36.0	265	1 UPPI_BACAN	Q81vc9 bacillus an
41	45	36.0	265	1 UPPI_BACHK	Q6hnb7 bacillus th
42	45	36.0	265	2 Q63G06	Q63G06 bacillus ce
43	45	36.0	285	2 Q9PB06	Q9PB06 xylella fas
44	45	36.0	356	1 YOP3 CAEEL	Q22695 caenorhabdi
45	45	36.0	420	2 Q22345	Q22345 zea mays (m

ALIGNMENTS

RESULT 1

Q9H7S7 PRELIMINARY; PRT; 1148 AA.

AC Q9H7S7; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishii T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fuji A., Hara H., Tanase T., Nomura Y.,
RA Toguya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamashita K., Watanabe K., Taniguchi A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujikawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togoishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RT Nat. Genet. 36:40-45 (2004).
RL EMBL; AK024365; BAB14899.1; -.
DR HSSP; Q9DHI1; 11VZ.
DR InterPro; IPR000082; SEA.
DR pfam; PF01390; SEA; 7.
DR PROSITE; PS50024; SEA; 3.
SQ SEQUENCE 1148 AA; 127957 MW; 3861B0D5EFPD8ABC CRC64;

```

Query Match          100.0%; Score 125; DB 2; Length 1148;
Best Local Similarity 100.0%; Pred. No. 4,2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SVPNRHTGVDSLGNFSPPLARRV 23
Db      1021 SVPNRHTGVDSLGNFSPPLARRV 1043

RESULT 2
Q96RK2      PRELIMINARY;      PRT; 6995 AA.
AC      O96RK2.
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Mucin 16 (Fragment).
GN      Name=MUC16.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200;
Yin B.W., Lloyd K.O.;
RA      Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF361486; AK74120.3; -.
DR      HSSP; Q9D1H1; 11VZ.
DR      InterPro; IPR000194; ATPase_a/bcentre.
DR      InterPro; IPR000082; SEA.
DR      Pfam; PF01390; SEA; 20.
DR      SMART; SM00200; SEA; 10.
DR      PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR      PROSITE; PS50024; SEA; 6.
RT      NON TER
SQ      SEQUENCE 6995 AA; 744958 MW; 80C797DBDF33A2B CRC64;

Query Match          100.0%; Score 125; DB 2; Length 6995;
Best Local Similarity 100.0%; Pred. No. 3,1e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SVPNRHTGVDSLGNFSPPLARRV 23
Db      6868 SVPNRHTGVDSLGNFSPPLARRV 6890

RESULT 3
Q9WK17      PRELIMINARY;      PRT; 22152 AA.
AC      Q9WK17.
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Ovarian cancer related tumor marker CA125.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RA      York L.;
RA      "The CA 125 gene: an extracellular superstructure dominated by repeat
RT      sequences.";
RL      Tumour Biol. 22:348-366(2001).

```

```

RN      [2]
RP      SEQUENCE FROM N.A.
RA      O'Brien T.J., Underwood L.J., Beard J.B.;
RA      Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF414442; AA65133.2; -.
DR      Genew; HGNC:15582; MUC16.
DR      Pfam; PF01390; SEA; 51.
DR      SMART; SM00200; SEA; 23.
DR      PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_2.
DR      PROSITE; PS50024; SEA; 11.
SQ      SEQUENCE 22152 AA; 2352668 MW; B3E7BDF1997A440 CRC64;

Query Match          100.0%; Score 125; DB 2; Length 22152;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SVPNRHTGVDSLGNFSPPLARRV 23
Db      22025 SVPNRHTGVDSLGNFSPPLARRV 22047

RESULT 4
Q9D1H1      PRELIMINARY;      PRT; 258 AA.
AC      Q9D1H1.
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE      library, clone:1110008114 product:hypothetical SEA domain containing
DE      protein, full insert sequence.
GN      Name=1110008114Rik;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J; TISSUE=whole body;
RA      MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA      Carninci P., Hayashizaki Y.;
RA      Carninci P., Hayashizaki Y.;
RL      "High-efficiency full-length cDNA cloning.";
RT      Meth. Enzymol. 303:19-44(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J; TISSUE=whole body;
RA      MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J; TISSUE=whole body;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J; TISSUE=whole body;
RA      MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA      Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA      Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT      "Normalization and subtraction of cap-trapper-selected cDNAs to
RT      prepare full-length cDNA libraries for rapid discovery of new genes.";
RL      Genome Res. 10:1617-1630(2000).
RN      [5]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6J; TISSUE=whole body;
RA      MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA      Shibata K., Itoh M., Aizawa K., Nagaoke S., Sasaki N., Carninci P.,
RA      Komno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

```

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashikawa K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayaishi Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-Format
 sequencing pipeline with 384 multicapillary sequencer.",
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=whole body;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carlinici P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanezaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK003577; BAB22869.1; -.
 DR PDB: 1IVZ; NMR: A=60-190.
 DR MGD: MGI:1920982; 111008114RIK.
 DR InterPro: IPR000082; SEA.
 DR Pfam: PF01390; SEA: 1.
 DR PROSITE: PS50024; SEA: 1.
 DR KW Hypothetical protein.
 SQ SEQUENCE 258 AA; 29425 MW; B64D9B63394D84E7 CRC64;

Query Match 78.4%; Score 98; DB 2; Length 258;
 Best Local Similarity 90.0%; Pred. No. 1.9e-07; Indels 0; Gaps 0;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 NRHHGTGVDSLCNFSPPLARV 23
 Db 134 NNNHTGVDSLCNFSPPLARR 153

RESULT 5
 ID Q98C55 PRELIMINARY; PRT; 141 AA.
 AC Q98C55;
 DT 01-OCT-2001 (T-EMBLrel. 18, Created)
 DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE M115290 protein.
 GN OrderedLocNames=M115290;
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=210829310; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.",
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003006; BAB51766.1; -.
 KW Complete proteome.
 SQ SEQUENCE 141 AA; 15437 MW; C43AC343668BEDAD CRC64;

Query Match 41.6%; Score 52; DB 2; Length 141;
 Best Local Similarity 40.0%; Pred. No. 2.9;
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 VPRHHGTGVDSLCNFSPPLAR 21
 Db 59 IPSKILPALFVCHFSPPAR 78

RESULT 6
 Y378 HALN1 STANDARD; PRT; 390 AA.
 ID Y378 HALN1
 AC Q9H570;
 DT 28-FEB-2003 (rel. 41, Created)
 DT 28-FEB-2003 (rel. 41, Last sequence update)
 DT 05-JUL-2004 (rel. 44, Last annotation update)
 DE Hypothetical UPF0204 protein Vng0378C.
 GN OrderedLocNames=VNG0378C;
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddock D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Beck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,
 RA Alam W., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hord L., Dassarma S.,
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- SIMILARITY: Belongs to the UPF0204 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AE004995; AAG18938.1; -.
 DR PIR: P84196; P84196.
 DR HAMAP: MF_00562; atypical; 1.
 DR InterPro: IPR007508; DUF516.
 DR Pfam: PF04414; DUF516; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 390 AA; 40610 MW; A2AF934F59498317 CRC64;

Query Match 41.6%; Score 52; DB 1; Length 390;
 Best Local Similarity 56.2%; Pred. No. 8.7;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 PNRHHGTGVDSLCNFSP 18
 Db 41 PNAHVAVDSLSAP 56

RESULT 7
 HEX_ADEG1 STANDARD; PRT; 942 AA.
 ID HEX_ADEG1
 AC P42511; 064758;
 DT 01-NOV-1995 (rel. 32, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 05-JUL-2004 (rel. 44, Last annotation update)
 DE Hexon protein (Late protein 2).
 GN Name=PII;
 OS Avian adenovirus gall (strain Phe1ps) (Powl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
 OX NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=96186720; PubMed=8627769;
RA Chioocca S., Kurzbauner R., Schaffner G., Baker A., Mautner V.,
RA Cocten M.;
RT "The complete DNA sequence and genomic organization of the avian
RT adenovirus CEO.";
RL J. Virol. 70:2939-2949(1996).
RN
RP SEQUENCE OF 913-942 FROM N.A.
RX MEDLINE=93362429; PubMed=8395124;
RA Cai F., Weber J.M.;
RT "Organization of the avian adenovirus genome and the structure of its
RT endoproteidase.";
RL Virology 196:358-362(1993).
CC -1- FUNCTION: This protein is one of the structural proteins in the
CC viral coat and is synthesized during late infection (By
CC similarity).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U46933; AAC54912.1; -.
DR EMBL: L13161; AAA51401.1; -.
DR HSSP: P03277; IDHX.
DR InterPro: IPR000736; Adeno_hexon.
DR Pfam: PF01065; Adeno_hexon; 1.
DR Pfam: PF03678; Adeno_hexon_C; 1.
DR ProDom: PD002815; Adeno_hexon; 1.
DR Coac protein; Hexon protein; Late protein.
SQ SEQUENCE 942 AA; 106709 MW; 7F4CE8D3F17D051B CRC64;

Query Match 40.8%; Score 51; DB 1; Length 942;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLGNFSPRLARR 22
Db 757 PDRHFHYDFLRNFDPMGRQ 776

RESULT 8
ID 096585 PRELIMINARY; PRT; 942 AA.
AC 096585;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Viral structural protein.
GN Name=hexon;
OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=FAV1;
RX MEDLINE=97049067; PubMed=8893797;
RA Akopian T.A., Doronin K.K., Karpov V.A., Naroditsky B.S.;
RT "Sequence of the avian adenovirus FAV1 (CELO) DNA encoding the hexon-
RT associated protein pIV and hexon.";
RL Arch. Virol. 141:1759-1765(1996).
DR HSBP: Z67970; CAA91908.1; -.
DR HSBP: P04133; IP30.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000736; Adeno_hexon.
DR InterPro: IPR001680; WD40.
DR Pfam: PF01065; Adeno_hexon; 1.
DR Pfam: PF03678; Adeno_hexon_C; 1.
DR ProDom: PD002815; Adeno_hexon; 1.
DR PROSITE: PS00678; WD REPEATS_1; UNKNOWN 1.
SQ SEQUENCE 942 AA; 106709 MW; 7F4CE8D3F17D051B CRC64;

```

```

DR ProDom: PD002815; Adeno_hexon; 1.
DR PROSITE: PS00678; WD REPEATS_1; UNKNOWN 1.
SQ SEQUENCE 942 AA; 106695 MW; 903732D3F5E6DC6F CRC64;

Query Match 40.8%; Score 51; DB 2; Length 942;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLGNFSPRLARR 22
Db 757 PDRHFHYDFLRNFDPMGRQ 776

RESULT 9
ID 07MSF4 PRELIMINARY; PRT; 942 AA.
AC 07MSF4;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE Hexon.
OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22935380; PubMed=14573794; DOI=10.1099/vir.0.19497-0;
RA Davidson A.J., Benko M., Harrach B.;
RT "Genetic content and evolution of adenoviruses.";
RL J. Gen. Virol. 84:2895-2908(2003)
CC -1- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/genbank/DBJ third party annotation (TPA) entry.
DR EMBL: BK001452; DAA01617.1; -.
DR HSSP: P04133; IP30.
DR GO: GO:0019028; C:structural molecule activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000736; Adeno_hexon.
DR InterPro: IPR001680; WD40.
DR Pfam: PF01065; Adeno_hexon; 1.
DR Pfam: PF03678; Adeno_hexon_C; 1.
DR ProDom: PD002815; Adeno_hexon; 1.
DR PROSITE: PS00678; WD REPEATS_1; UNKNOWN 1.
SQ SEQUENCE 942 AA; 106709 MW; 7F4CE8D3F17D051B CRC64;

Query Match 40.8%; Score 51; DB 2; Length 942;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 PNRHHTGVDSLGNFSPRLARR 22
Db 757 PDRHFHYDFLRNFDPMGRQ 776

RESULT 10
ID 06LAN2 PRELIMINARY; PRT; 785 AA.
AC 06LAN2;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein P0473H02.11.
GN Name=P0473H02.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacoidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Heng Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Lwu H.-L.,

```

us-10-612-090-17.rup

DR EMBL; AB009688; BAA33384.1; -
DR EMBL; AB009689; BAA33385.1; -
DR EMBL; AB009689; BAA33386.1; -
DR EMBL; AF035161; AAF13302.1; -
DR MGI; 1278342; Lect2.
DR InterPro; IPR008663; LECT2.
DR Pfam; PF05429; LECT2; 1.
KW Alternative splicing; Chemotaxis; Signal.
FT SIGNAL
FT CHAIN 1
FT VARSPUBLIC 98 151
SQ VARIANT 129 129 16405 MW; 18AFA44046B7ABE CRC64;

Query Match 39.6%; Score 49.5; DB 1; Length 151;
Best Local Similarity 37.5%; Pred. No. 7.8;
Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 1;

QY 5 RHHTGVDSLCN-----FSLPARV 23
Db 51 RHHGVDVLCDSGSVYVAAPTGR 74

RESULT 12
ORXK181 PRELIMINARY; PRT; 151 AA.
ID ORXK181
AC ORXK181;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Leukocyte cell-derived chemotaxin 2.
GN Name=lect2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Yoshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McClean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schercher A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX Strausberg R.;
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027753; AA027753.1; -
DR MGI; MG11278342; Lect2.
DR GO; GO:0005615; Cytoextracellular space; VAS.
DR InterPro; IPR008663; LECT2.
Pfam; PF05429; LECT2; 1.

SQ SEQUENCE 151 AA; 16439 MW; B7AF445C2B017FE2 CRC64;
 Query Match 39.6%; Score 49.5; DB 2; Length 151;
 Best Local Similarity 37.5%; Pred. No. 7.8;
 Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 1;
 QY 5 RHHTGVDSLNCN----FSPPLARV 23
 DB 51 RHHPGVDLCSGDSVYAPFTGKI 74
 RESULT 13
 Q9XDT0 PRELIMINARY; PRT; 207 AA.
 AC Q9XDT0;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Neud.
 GN Name=neud;
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99395021; PubMed=10464185;
 RA Yamamoto S., Miyake K., Koike Y., Watanabe M., Machida Y., Ohta M.,
 RA Iijima S.;
 RA "Molecular characterization of type-specific capsular polysaccharide
 RT biosynthesis genes of Streptococcus agalactiae type Ia."
 RL J. Bacteriol. 181:5176-5184(1999).
 DR EMBL; AB028896; BAA82289.2; -
 DR InterPro: IPR011004; Trimer LpxA like.
 SQ SEQUENCE 207 AA; 22397 MW; D3329979F9E68C53 CRC64;
 Query Match 39.2%; Score 49; DB 2; Length 207;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 6 HHTGVDSLNCNPSPLA 20
 DB 138 HHTVESHCHNIAFNA 152
 RESULT 14
 Q9RPB8 PRELIMINARY; PRT; 209 AA.
 AC Q9RPB8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Neud.
 GN Name=neud;
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COH1;
 RC MEDLINE=96228704; PubMed=8830246;
 RA Haft R.F., Weesels M.R., Mebane M.F., Conaty N., Rubens C.E.;
 RA "Characterization of cpsF and its product CMP-N-acetylneuraminic acid
 RT synthetase, a group B streptococcal enzyme that can function in KI
 RL capsular polysaccharide biosynthesis in Escherichia coli."
 RL Mol. Microbiol. 19:555-563(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=COH1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20372630; PubMed=10913080;
 RX DOI=10.1128/DB.182.16.4466-4477.2000;
 RA Chaffin D.O., Beres S.B., Yim H.H., Rubens C.E.;

RT "The serotype of type Ia and III group B streptococci is determined by
 RT the polymerase gene within the polyclonistic capsule operon."
 RL J. Bacteriol. 182:4466-4477(2000).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NT6;
 RA Mckinnon K., Chaffin D.O., Rubens C.E.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CNCTC 1/82;
 RA Mckinnon K., Chaffin D.O., Rubens C.E.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CNCTC 1/82;
 RA Mckinnon K., Chaffin D.O., Rubens C.E.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=SMU014, and 7271;
 RA Cieslewicz M.J., Glusman G., Chaffin D., Kasper D., Madan A.,
 RA Rodrigues S., Fahey J., Weesels M.R., Rubens C.E.;
 RT "Evolution of group B Streptococcus capsular polysaccharides."
 RT Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF163833; AAD53076.1; -
 DR EMBL; AF163833; AAD53076.1; -
 DR EMBL; AF349539; AAK29663.1; -
 DR EMBL; AF349539; AAK29663.1; -
 DR EMBL; AF355776; AAK3617.1; -
 DR EMBL; AY375362; AAR29918.1; -
 DR EMBL; AY375363; AAR29920.1; -
 DR EMBL; AY376403; AAR29958.1; -
 DR PIR; T44653; T44653.
 DR InterPro: IPR011004; Trimer LpxA like.
 SQ SEQUENCE 209 AA; 22578 MW; BD07433783C952E33 CRC64;
 Query Match 39.2%; Score 49; DB 2; Length 209;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 6 HHTGVDSLNCNPSPLA 20
 DB 138 HHTVESHCHNIAFNA 152
 RESULT 15
 Q8DZES PRELIMINARY; PRT; 209 AA.
 AC Q8DZES;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Neud protein.
 GN Name=neud; OrderedLocustNames=SA61159;
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RC MEDLINE=22229388; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Tetteelin H., Maignani V., Cieslewicz M.J., Eileen J.A., Peterson S.N.,
 RA Weesels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA Debay R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Knout H.M., Mulligan S.,
 RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rapoport R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae."

RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL; AE014244; AAN00041.1; -.
 DR PIR; T44653; T44653.
 DR TIGR; SAG1159; -.
 DR InterPro; IPR011004; Trimer_LpxA_like.
 KW Complete proteome.
 SQ SEQUENCE 209 AA; 22578 MW; B007433783C852E3 CRC64;

Query Match 39.2%; Score 49; DB 2; Length 209;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 HHTGVDSLGNFSPPLA 20
 |||||:|||||
 Db 138 HHTVESHCHNAPNA 152

Search completed: October 27, 2005, 18:55:21
 Job time : 69.8605 secs

.....

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:31:33 ; Search time 96.893 Seconds
(without alignments)
111.765 Million cell updates/sec

Title: US-10-612-090-18

Perfect score: 141
Sequence: 1 DRVAIVEEFLRMTRNGTQLQNFTRDRSS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	141	100.0	28	8	ADJ66745	Adj66745 Human Muc
2	141	100.0	108	8	ADJ66729	Adj66729 Human Muc
3	141	100.0	108	8	ADJ66740	Adj66740 Human Muc
4	141	100.0	174	7	ADM50756	Adm50756 Human can
5	141	100.0	178	5	ABP30979	Abp30979 Predicted
6	141	100.0	178	7	ADA08642	Ada08642 Human O77
7	141	100.0	178	7	ADF08985	Adf08985 Secreted
8	141	100.0	233	5	ABP30978	Abp30978 Extracell
9	141	100.0	233	7	ADA08641	Ada08641 Human O77
10	141	100.0	233	7	ADF08984	Adf08984 Secreted
11	141	100.0	284	6	ABU54859	Abu54859 Human CA1
12	141	100.0	318	5	ABP31025	Abp31025 Amino aci
13	141	100.0	318	7	ADA08747	Ada08747 Human O77
14	141	100.0	318	7	ADF09090	Adf09090 Secreted
15	141	100.0	367	4	AAE12631	Aae12631 Human gen
16	141	100.0	438	3	AAE12555	Aae12555 Human ova
17	141	100.0	438	4	AAE12623	Aae12623 Human gen
18	141	100.0	438	5	ABP30973	Abp30973 Partial p
19	141	100.0	438	5	ABP30899	Abp30899 0772P c10
20	141	100.0	438	7	ADA08636	Ada08636 Human O77
21	141	100.0	438	7	ADA08543	Ada08543 Human ova
22	141	100.0	438	7	ADP08886	Adp08886 Secreted
23	141	100.0	438	7	ADP08979	Adp08979 Secreted
24	141	100.0	438	7	ADG46173	Adg46173 Human ova
25	141	100.0	439	6	ABU54709	Abu54709 Human CA1

26	141	100.0	526	4	AAE12634	Aae12634 Human gen
27	141	100.0	748	8	ADS94302	Ads94302 CA 125/07
28	141	100.0	809	8	ADS94303	Ads94303 CA 125/07
29	141	100.0	833	3	AAE12554	Aae12554 Human ova
30	141	100.0	833	5	ABP30898	Abp30898 0772P c10
31	141	100.0	833	7	ADA08542	Ada08542 Human ova
32	141	100.0	833	7	ADP08885	Adp08885 Secreted
33	141	100.0	833	7	ADG46172	Adg46172 Human ova
34	141	100.0	914	3	AAE12552	Aae12552 Human ova
35	141	100.0	914	4	AAE99203	Aae99203 Human ova
36	141	100.0	914	5	ABP30868	Abp30868 Hypotheti
37	141	100.0	914	5	ABP30896	Abp30896 0772P pro
38	141	100.0	914	7	ADA08631	Ada08631 Human O77
39	141	100.0	914	7	ADA08465	Ada08465 Human ova
40	141	100.0	914	7	ADP08974	Adp08974 Secreted
41	141	100.0	914	7	ADP08808	Adp08808 Secreted
42	141	100.0	914	7	ADG46095	Adg46095 Human ova
43	141	100.0	914	8	ADN40451	Adn40451 Human bre
44	141	100.0	1148	4	AAE95836	Aae95836 Human pro
45	141	100.0	1148	4	ABE50283	AbE50283 HOST-1 ov

ALIGNMENTS

RESULT 1
ADJ66745
ID ADJ66745 standard; peptide; 28 AA.
XX
XX ADJ66745;
XX
XX 06-MAY-2004 (first entry)
XX
XX Human Muc1/Muc16 protein-related peptide SeqID18.
XX
XX monoclonal antibody; epitope; non-shed extracellular portion;
XX shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
XX breast cancer; ovarian cancer.
XX
XX Homo sapiens.
XX
XX WO2004005470-A2.
XX
XX PD 15-JAN-2004.
XX
XX PF 03-JUL-2003; 2003WO-US020907.
XX
XX PR 03-JUL-2002; 2002US-0393094P.
XX
XX (IMMU-) IMMUNOGEN INC.
XX
XX Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K;
XX Vater CA;
XX
XX WPI; 2004-091350/09.
XX
XX PT New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,
XX treating or monitoring malignancies, such as breast or ovarian cancer.
XX
XX PS Claim 14; SEQ ID NO 18; 113pp; English.
XX
XX This invention relates to a novel isolated monoclonal antibody that
XX specifically binds to an epitope of a non-shed extracellular portion of a
XX shed antigen or of human Muc1 or Muc16 protein. The invention may be
XX useful for the development of compounds with a cytostatic or for gene
XX therapy. The composition and methods are useful in diagnosing, treating
XX or monitoring malignancies, such as breast or ovarian cancer. The present
XX sequence is that of a peptide which was used for raising antibodies to
XX the extracellular, non-shed region of Muc16 and which is claimed in the
XX specification.
XX
XX Sequence 28 AA:

Query Match 100.0%; Score 141; DB 8; Length 28;
 Best Local Similarity 100.0%; Pred. No. 7.3e-15;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DRVAIYEEFLMRTNGTOLQNTLDRSS 28
 1 DRVAIYEEFLMRTNGTOLQNTLDRSS 28

RESULT 2

ADJ66729
 ID ADJ66729 standard; protein; 108 AA.

AC ADJ66729;

DT 06-MAY-2004 (first entry)

XX Human Muc16 epitope amino acid sequence.

XX monoclonal antibody; epitope; non-shed extracellular portion;

KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;

KW breast cancer; ovarian cancer.

XX Homo sapiens.

PN WO2004005470-A2.

PD 15-JAN-2004.

XX 03-JUL-2003; 2003WO-US020907.

PR 03-JUL-2002; 2002US-0393094P.

XX (IMMU-) IMMUNOGEN INC.

PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K,

PT Vater CA;

XX WPI; 2004-091350/09.

XX New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,

PT treating or monitoring malignancies, such as breast or ovarian cancer.

XX Claim 13; SEQ ID NO 2; 113pp; English.

XX This invention relates to a novel isolated monoclonal antibody that

CC specifically binds to an epitope of a non-shed extracellular portion of a

CC shed antigen or of human Muc1 or Muc16 protein. The invention may be

CC useful for the development of compounds with a cytostatic or for gene

CC therapy. The composition and methods are useful in diagnosing, treating

CC or monitoring malignancies, such as breast or ovarian cancer. The present

CC sequence is that of a Muc epitope of the invention.

XX Sequence 108 AA;

XX Query Match 100.0%; Score 141; DB 8; Length 108;

XX Best Local Similarity 100.0%; Pred. No. 3.6e-14;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DRVAIYEEFLMRTNGTOLQNTLDRSS 28
 60 DRVAIYEEFLMRTNGTOLQNTLDRSS 87

RESULT 3

ADJ66740
 ID ADJ66740 standard; protein; 108 AA.

AC ADJ66740;

DT 06-MAY-2004 (first entry)

XX Human Muc16 GST fusion protein amino acid sequence SeqID13.

XX monoclonal antibody; epitope; non-shed extracellular portion;
 KW shed antigen; human; Muc1; Muc16; cytostatic; gene therapy; malignancy;
 KW breast cancer; ovarian cancer.

XX Homo sapiens.

PN WO2004005470-A2.

PD 15-JAN-2004.

XX 03-JUL-2003; 2003WO-US020907.

PR 03-JUL-2002; 2002US-0393094P.

XX (IMMU-) IMMUNOGEN INC.

PI Payne G, Chittenden T, Goldmakher V, Chun P, Sneider-Mulready K,

PT Vater CA;

XX WPI; 2004-091350/09.

XX New antibody to non-shed human Muc1 or Muc16, useful for diagnosing,

PT treating or monitoring malignancies, such as breast or ovarian cancer.

XX Disclosure; SEQ ID NO 13; 113pp; English.

XX This invention relates to a novel isolated monoclonal antibody that

CC specifically binds to an epitope of a non-shed extracellular portion of a

CC shed antigen or of human Muc1 or Muc16 protein. The invention may be

CC useful for the development of compounds with a cytostatic or for gene

CC therapy. The composition and methods are useful in diagnosing, treating

CC or monitoring malignancies, such as breast or ovarian cancer. The present

CC sequence is that of the human Muc16 GST fusion protein which was used in

CC the exemplification of the invention.

XX Sequence 108 AA;

XX Query Match 100.0%; Score 141; DB 8; Length 108;

XX Best Local Similarity 100.0%; Pred. No. 3.6e-14;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DRVAIYEEFLMRTNGTOLQNTLDRSS 28
 60 DRVAIYEEFLMRTNGTOLQNTLDRSS 87

ADJ30756
 ID ADJ30756 standard; protein; 174 AA.

XX ADM30756;

XX 20-MAY-2004 (first entry)

XX Human cancer linked protein referred to as 192292 SeqID 4.

XX human; cancer-related gene; screening assay; immuno-conjugate;

XX cytostatic; immunostimulant; cancer; ovarian cancer.

XX Homo sapiens.

PN WO2003075854-A2.

PD 18-SEP-2003.

XX 07-MAR-2003; 2003WO-US007147.

XX 07-MAR-2002; 2002US-0362527P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Ebner R, Strovel JW;

XX WPI; 2003-748320/70.
DR N-PSDB; ADM30755.
XX
PT Identifying an agent that modulates the activity of any of three specific
PT cancer-related genes, potentially useful in treating (ovarian) cancer,
PT comprises detecting a difference in expression of the gene in the
PT presence of the agent.
XX
PS Claim 11; SEQ ID NO 4; 57bp; English.
XX
CC This invention relates to a novel method of identifying an agent that
CC modulates the activity of a cancer-related gene. Specifically, it refers
CC to a screening assay to identify potential antitumour agents, as well as
CC methods to assess the cancerous state of a cell. The present invention
CC describes antibodies against the expression products of these cancer-
CC related genes that are capable of targeting a cancerous cell in vivo.
CC Furthermore it provides immuno-conjugates containing such antibodies,
CC which can be used to deliver target therapeutics thereto. Accordingly,
CC these compositions that exhibit cytostatic and immunostimulant activities
CC are useful for preventing or treating cancer in humans and in particular
CC the treatment of ovarian cancer. This polypeptide sequence is a protein
CC encoded by a human cancer-linked gene of the invention.
SQ
Sequence 174 AA;
Query Match 100.0%; Score 141; DB 7; Length 174;
Best Local Similarity 100.0%; Pred. No. 6.3e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRVAIYEEFLRMTRNGTQLQNFLLDRSS 28
DB 70 DRVAIYEEFLRMTRNGTQLQNFLLDRSS 97
RESULT 5
ABP30979
ID ABP30979 standard; protein; 178 AA.
XX
XX ABP30979;
XX
XX 02-JUL-2002 (first entry)
XX
XX Predicted extracellular domain of 0772P.
XX
XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
XX
XX Homo sapiens.
XX
XX WO200206317-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001WO-US022635.
XX
XX 17-JUL-2000; 2000US-00617747.
XX
XX 10-AUG-2000; 2000US-00636801.
XX
XX 20-SEP-2000; 2000US-00667857.
XX
XX 04-APR-2001; 2001US-00827271.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX
XX Reed SG, Vedvyck TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2002-164781/21.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
XX protein or its variants, useful for stimulating an immune response in a
XX patient and treating ovarian cancer.
XX
XX Example 13; Page 361-362; 408bp; English.

XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX
SQ
Sequence 178 AA;
Query Match 100.0%; Score 141; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRVAIYEEFLRMTRNGTQLQNFLLDRSS 28
DB 129 DRVAIYEEFLRMTRNGTQLQNFLLDRSS 156
RESULT 6
ADA08642
ID ADA08642 standard; protein; 178 AA.
XX
XX ADA08642;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human 0772P partial protein #12.
XX
XX human; gene therapy; ovarian cancer; cancer.
XX
XX Homo sapiens.
XX
XX OS
XX PN US2003091580-A1.
XX
XX 15-MAY-2003.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (MITC/) MITCHAM J L.
XX
XX (KING/) KING G E.
XX
XX (ALGA/) ALGATE P A.
XX
XX (FLIN/) FLING S P.
XX
XX (RETT/) RETTER M W.
XX
XX (FANG/) FANGER G R.
XX
XX (REED/) REED S G.
XX
XX (VEDV/) VEDDYCK T S.
XX
XX (CART/) CARTER D.
XX
XX (HILL/) HILL P.
XX
XX (ALBO/) ALBONE E.
XX
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX
XX Reed SG, Vedvyck TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
XX
XX New isolated 0772P polypeptides and polynucleotides, useful in gene
XX therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX cancer.
XX
XX Disclosure; Page 28; 371bp; English.
XX
XX The invention relates to an isolated 0772P polypeptide, which has the
XX structure fully defined in the specification. The composition containing
XX the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX or antigen presenting cells are useful for stimulating an immune response
XX and treating ovarian cancer. Detecting the presence of the
XX polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX carcinoma cDNAs and protein cDNAs were identified using microarray
XX technology. The present sequence represents a human ovarian carcinoma

CC antigen.
 XX
 SQ Sequence 178 AA;
 Query Match 100.0%; Score 141; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 6,4e-14;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DRVAIYEEFLMTRNGTQLQNFLLDRSS 28
 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156
 Db 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156
 RESULT 7
 ADF08985
 ID ADF08985 standard; protein; 178 AA.
 AC ADF08985;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Secreted ovarian carcinoma antigen seqid 489.
 XX
 KW gene therapy; protein therapy; vaccine; antibody inhibition;
 KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
 KW secreted ovarian carcinoma antigen.
 XX
 OS Homo sapiens.
 XX
 PN US2003124140-A1.
 PD 03-JUL-2003.
 XX
 PF 17-JUL-2002; 2002US-00198053.
 XX
 PR 17-DEC-1998; 98US-00215681.
 PR 17-DEC-1998; 98US-00216003.
 PR 23-JUN-1999; 99US-00338933.
 PR 24-SEP-1999; 99US-00404879.
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 PR 17-JUL-2001; 2001US-00907969.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Bangur CS, Retter MW, Fanger GR, Hill P;
 DR WPI; 2003-897152/82.
 XX
 PT Oncogenic nucleic acids useful for the prevention, diagnosis and
 PT treatment of breast cancer.
 XX
 PS Example 13; SEQ ID NO 489; 399bp; English.
 XX
 CC The invention describes nucleic acids (I) and the polypeptides (II) they
 CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
 CC treating diseases related to their aberrant expression i.e. breast
 CC cancers. For example, (I) and (II) may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of (II) by expressing
 CC inactive proteins or to supplement the patients own production of (II).
 CC Additionally, (I) may be used to produce (II), by inserting (I) into a
 CC host cell and culturing the cell to express the protein (II). (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. The host cell may also be used as antigens in the production of
 CC antibodies against (II) and in assays to identify modulators of (II)'s
 CC expression and activity. The anti-(II) antibodies, agonists and as
 CC antagonists may be used to regulate expression and activity and as

CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
 CC immunosassay). This sequence represents a secreted ovarian carcinoma
 CC antigen.
 XX
 SQ Sequence 178 AA;
 Query Match 100.0%; Score 141; DB 7; Length 178;
 Best Local Similarity 100.0%; Pred. No. 6,4e-14;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DRVAIYEEFLMTRNGTQLQNFLLDRSS 28
 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156
 Db 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156
 RESULT 8
 ABP30978
 ID ABP30978 standard; protein; 233 AA.
 AC ABP30978;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Extracellular and transmembranal regions of 0772P.
 XX
 KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200206317-A2.
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-US022635.
 XX
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-00667857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 XX
 DR WPI; 2002-164781/21.
 XX
 PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 PS Example 13; Page 361; 408bp; English.
 XX
 CC This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents protein
 CC related to the invention
 CC
 SQ Sequence 233 AA;
 Query Match 100.0%; Score 141; DB 5; Length 233;
 Best Local Similarity 100.0%; Pred. No. 8,8e-14;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DRVAIYEEFLMTRNGTQLQNFLLDRSS 28
 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156
 Db 129 DRVAIYEEFLMTRNGTQLQNFLLDRSS 156

RESULT 9
ADA08641
ADA08641 standard; protein; 233 AA.
XX
AC ADA08641;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human O772P partial protein #11.
XX
KW human; gene therapy; ovarian cancer; cancer.
XX
OS Homo sapiens.
XX
PN US2003091580-A1.
XX
PD 15-MAY-2003.
XX
PF 17-JUL-2001; 2001US-00907969.
XX
PR 18-JUN-2001; 2001US-0084441.
XX
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2003-532352/50.
XX
PT New isolated O772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT cancer.
XX
PS Example 13; Page 28; 371pp; English.
XX
CC The invention relates to an isolated O772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen.
XX
SQ Sequence 233 AA;
XX
Query Match 100.0%; Score 141; DB 7; Length 233;
Best Local Similarity 100.0%; Pred. No. 8.8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRVAIYEEFLRMTRNGTOLONFTLDRSS 28
DB 129 DRVAIYEEFLRMTRNGTOLONFTLDRSS 156

XX
DE Secreted ovarian carcinoma antigen segid 488.
XX
KW gene therapy; protein therapy; vaccine; antibody inhibition;
KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW secreted ovarian carcinoma antigen.
XX
OS Homo sapiens.
XX
PN US2003124140-A1.
XX
PD 03-JUL-2003.
XX
PF 17-JUL-2002; 2002US-00198053.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-0084441.
PR 17-JUL-2001; 2001US-00907969.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2003-897152/82.
XX
PT Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
PS Example 13; SEQ ID NO 488; 399pp; English.
XX
CC The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patient's own production of (II).
CC Additionally, (II) may be used to produce (II), by inserting (II) into a
CC host cell and culturing the cell to express the protein (II). (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The host cell may also be used as antigens in the production of
CC antibodies against (II) and in assays to identify modulators of (II)'s
CC expression and activity. The anti-(II) antibodies, agonists and
CC antagonists may be used to regulate expression and activity and as
CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC immunoassay). This sequence represents a secreted ovarian carcinoma
CC antigen.
XX
SQ Sequence 233 AA;
XX
Query Match 100.0%; Score 141; DB 7; Length 233;
Best Local Similarity 100.0%; Pred. No. 8.8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DRVAIYEEFLRMTRNGTOLONFTLDRSS 28
DB 129 DRVAIYEEFLRMTRNGTOLONFTLDRSS 156

RESULT 10
ADF08984
ADF08984 standard; peptide; 233 AA.
XX
AC ADF08984;
XX
DT 12-FEB-2004 (first entry)

RESULT 11
ABU54859
ABU54859 standard; protein; 284 AA.
XX
AC ABU54859;

XX 12-MAR-2003 (first entry)
 DT Human CA125 carboxy terminal domain.
 DE
 XX
 KW Human; CA125; protein repeat; chromosome 19q 13.2; amino terminal domain;
 XX amino terminal extension; carboxy terminal domain; vaccine; cancer;
 KW ovarian cancer; carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN MO200283866-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US011734.
 XX
 PR 17-APR-2001; 2001US-0284175P.
 XX 19-JUN-2001; 2001US-0299380P.
 PR 27-SEP-2001; 2001US-00965738.
 PR 21-DEC-2001; 2001US-0345180P.
 XX
 PA (UYAR-) UNITV ARKANSAS.
 XX
 PI O'brien T, Beard J, Underwood L;
 DR WPI; 2003-093013/08.
 XX
 PT New CA125 molecules, useful as a gold standard for detecting and
 PT monitoring the presence of CA125 antigen which can be used for
 PT diagnosing, monitoring or treating patients with cancer or for developing
 PT vaccine against cancer.
 XX
 PS Claim 1; Fig 9b; 694pp; English.
 XX
 CC The invention relates to a CA125 protein comprising: (a) an extracellular
 CC amino terminal domain; (b) an amino terminal extension; (c) a multiple
 CC repeat domain; and (d) a carboxy terminal domain. The extracellular amino
 CC terminal domain comprises 5 genomic exons, the amino terminal extension
 CC comprises 4 genomic exons, each repeat unit comprises 5 genomic exons and
 CC the carboxy terminal domain comprises a transmembrane anchor with a short
 CC cytoplasmic domain, and further comprises 9 genomic exons. The gene for
 CC CA125 is located on human chromosome 19q 13.2. Also included are isolated
 CC CA125 repeat domains, nucleic acids (including variants, homologues and
 CC degenerate versions) encoding CA125 proteins or repeat units, a vector
 CC comprising the nucleic acid, a cultured cell comprising the vector, a
 CC method of expressing CA125 antigen in a cell, the amino acid sequences of
 CC the CA125 repeat units (or their variants, fragments or sequences 50%
 CC identical to them), a purified antibody that selectively binds to an
 CC epitope in the receptor-binding domain of CA125 protein, a diagnostic for
 CC detecting and monitoring the presence of CA125 antigen (comprising
 CC recombinant CA125 having at least one repeat unit of the CA125 repeat
 CC domain including epitope binding sites), a therapeutic vaccine to treat
 CC mammals with elevated CA125 antigen levels or at risk of developing a
 CC disease or disease recurrence associated with elevated CA125 antigen
 CC levels (comprising recombinant CA125 repeat domains including epitope
 CC binding sites) and an antisense oligonucleotide that inhibits the
 CC expression of CA125. The CA125 molecule, particularly the multiple repeat
 CC domains are useful as a gold standard for detecting and monitoring the
 CC presence of CA125 antigen, which can be used for diagnosing, monitoring
 CC or treating patients with ovarian cancer and other carcinomas where CA125
 CC is expressed. The molecules are also useful for developing a vaccine
 CC against cancer. The present sequence is a CA125 repeat protein
 XX
 SQ Sequence 284 AA;
 Query Match 100.0%; Score 141; DB 6; Length 284;
 Best Local Similarity 100.0%; Pred. No. 1,1e-13;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRVAIYEEFLMTRNGTQLQNFLLDRSS 28
 DB 180 DRVAIYEEFLMTRNGTQLQNFLLDRSS 207

RESULT 12
 ID ABP31025
 AC ABP31025 standard; protein; 318 AA.
 XX
 AC ABP31025;
 XX
 DT 02-JUL-2002 (first entry)
 DE
 XX
 DE Amino acid sequence of the 3' constant region of 0772P.
 XX
 KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN MO200206317-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-US022635.
 XX
 PR 17-JUL-2000; 2000US-00617747.
 PR 10-AUG-2000; 2000US-00636801.
 PR 20-SEP-2000; 2000US-0067857.
 PR 04-APR-2001; 2001US-00827271.
 PR 18-JUN-2001; 2001US-00884441.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham UL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 DR WPI; 2002-164781/21.
 XX
 PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
 PT protein or its variants, useful for stimulating an immune response in a
 PT patient and treating ovarian cancer.
 XX
 PS Claim 3; Page 399-400; 408pp; English.
 XX
 CC This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents protein
 CC related to the invention
 XX
 SQ Sequence 318 AA;
 Query Match 100.0%; Score 141; DB 5; Length 318;
 Best Local Similarity 100.0%; Pred. No. 1,3e-13;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRVAIYEEFLMTRNGTQLQNFLLDRSS 28
 DB 214 DRVAIYEEFLMTRNGTQLQNFLLDRSS 241
 RESULT 13
 ID ADA08747
 AC ADA08747 standard; protein; 318 AA.
 XX
 AC ADA08747;
 XX
 DT 06-NOV-2003 (first entry)
 DE
 XX
 DE Human O772P constant region consensus sequence #1.
 XX
 KW human; gene therapy; ovarian cancer; cancer.
 XX
 OS Homo sapiens.

XX US2003091580-A1.
PN
XX
XX 15-MAY-2003.
PD
XX
XX 17-JUL-2001; 2001US-00907969.
PF
XX
XX 18-JUN-2001; 2001US-00844441.
PR
XX
XX (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDY/) VEDYICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
XX New isolated 0772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX cancer.
XX
XX Claim 3; SEQ ID NO 594; 371pp; English.
XX
XX The invention relates to an isolated 0772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma CDNAs and protein CDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen.
XX
XX Sequence 318 AA;
SQ

Query Match 100.0%; Score 141; DB 7; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 214 DRVAIYEEFLRMTRNGTQLONFLLDRSS 241

RESULT 14
ADFO9090 ID ADF09090 standard; protein; 318 AA.
XX
XX ADF09090;
AC
XX 12-FEB-2004 (first entry)
DT
XX
XX Secreted ovarian carcinoma antigen seqid 594.
DE
XX
XX gene therapy; protein therapy; vaccine; antibody inhibition;
KM breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW secreted ovarian carcinoma antigen.
XX
XX Homo sapiens.
OS
XX US2003124140-A1.
PN
XX 03-JUL-2003.
PD
XX

PF 17-JUL-2002; 2002US-00198053.
XX
XX 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
XX
XX (CORI-) CORIXA CORP.
XX
XX Bangur CS, Retter MW, Fanger GR, Hill P;
PI WPI; 2003-897152/82.
XX
XX N-PSDB; ADF09064.
DR
XX
XX Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
XX Example 16; SEQ ID NO 594; 399pp; English.
XX
XX The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patients own production of (II).
CC Additionally, (I) may be used to produce (II), by inserting (I) into a
CC host cell and culturing the cell to express the protein (II). (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The host cell may also be used as antigens in the production of
CC antibodies against (II) and in assays to identify modulators of (II)'s
CC expression and activity. The anti-(II) antibodies, agonists and
CC antagonists may be used to regulate expression and activity and as
CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC immunoassay). This sequence represents a secreted ovarian carcinoma
CC antigen.
XX
XX Sequence 318 AA;
SQ

Query Match 100.0%; Score 141; DB 7; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 214 DRVAIYEEFLRMTRNGTQLONFLLDRSS 241

RESULT 15
AAE12631 ID AAE12631 standard; protein; 367 AA.
XX
XX AAE12631;
AC
XX 03-JAN-2002 (first entry)
DT
XX
XX Human gene 1 encoded secreted protein fragment, SEQ ID NO: 27.
DE
XX
XX Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;
KM respiratory system disorder; asthma; haematopoietic disorder; skin aging;
KW immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
KW rheumatoid arthritis; inflammation; neurological disorder; septic shock;
KW Alzheimer's disease; Parkinson's disease; diabetes; angiogenic disorder;
KW atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
KW epithelial cell proliferation; transplantation; chemotaxis; infection;
KW

KM food additive; wound healing; endocrine disorder; kidney disorder;
 KW gene therapy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN MO200170804-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 16-MAR-2001; 2001WO-US008585.
 XX
 PR 17-MAR-2000; 2000US-0190076P.
 PR 23-AUG-2000; 2000US-0227009P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2001-639119/73.
 XX
 PT Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful for
 PT preventing, diagnosing and/or treating ovarian cancer and ovarian cancer
 PT metastases.
 XX
 PS Disclosure; Page 9; 427pp; English.
 XX
 CC AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian and
 CC ovarian cancer associated protein (collectively known as ovarian cancer
 CC antigens) genes, and AAE12623-AAE12629 represent the proteins they
 CC encode. AAE12630-AAE12638 represent human ovarian cancer antigenic
 CC fragments or variants. Ovarian cancer antigens and their corresponding
 CC DNAs are used in the prevention, diagnosis and treatment of diseases
 CC associated with their inappropriate expression. These disorders include
 CC proliferative disorders, cancer, tumours, respiratory system disorders,
 CC asthma, haematopoietic disorders, diseases of the immune system, AIDS,
 CC skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
 CC inflammation, allergies, neurological disorders (e.g., Alzheimer's
 CC disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotonic disorders, kidney disorders,
 CC gastrointestinal disorders, endocrine disorders and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, to identify their cognate ligands or binding
 CC partners, in chemotaxis and can be used as a food additive. Antibodies
 CC specific for a protein of the invention can be used in alleviating
 CC symptoms associated with the disorders mentioned above and in diagnostic
 CC immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
 CC the invention is used in gene therapy. The present sequence represents a
 CC human ovarian cancer antigenic fragment of the invention
 XX
 SQ Sequence 367 AA;
 XX
 Query Match 100.0%; Score 141; DB 4; Length 367;
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DRVAIYEEFLRMTRNGTOLQNFLLDRSS 28
 |||||
 DB 334 DRVAIYEEFLRMTRNGTOLQNFLLDRSS 361

Search completed: October 27, 2005, 18:44:25
 Job time : 96.893 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:03:42 ; Search time 24.3535 Seconds
(without alignments)
85.826 Million cell updates/sec

Title: US-10-612-090-18

Perfect score: 141
Sequence: 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	100.0	438	4 US-09-404-879A-390	Sequence 390, App
2	141	100.0	438	4 US-09-667-857-390	Sequence 390, App
3	141	100.0	833	4 US-09-404-879A-389	Sequence 389, App
4	141	100.0	833	4 US-09-667-857-389	Sequence 389, App
5	141	100.0	914	4 US-09-404-879A-312	Sequence 312, App
6	141	100.0	914	4 US-09-338-933-312	Sequence 312, App
7	141	100.0	914	4 US-09-667-857-312	Sequence 312, App
8	56	39.7	772	4 US-09-404-879A-388	Sequence 388, App
9	56	39.7	772	4 US-09-667-857-388	Sequence 388, App
10	48.5	34.4	263	4 US-09-710-279-332	Sequence 332, App
11	48.5	34.4	274	3 US-09-134-001C-3921	Sequence 3921, App
12	47	33.3	514	3 US-09-066-047-3	Sequence 3, App1
13	47	33.3	2088	4 US-09-802-540-12906	Sequence 12906, A
14	46.5	33.0	693	4 US-09-818-780-68	Sequence 68, App1
15	46	32.6	115	4 US-09-107-532A-5481	Sequence 5481, App
16	45.5	32.3	433	4 US-09-328-352-75498	Sequence 7548, App
17	45	31.9	245	4 US-09-270-767-45297	Sequence 45297, A
18	45	31.9	255	4 US-09-270-767-32769	Sequence 32769, A
19	45	31.9	255	4 US-09-270-767-47986	Sequence 47986, A
20	45	31.9	275	4 US-09-338-092-348	Sequence 47986, A
21	44.5	31.6	145	3 US-08-946-329A-46	Sequence 46, App1
22	44.5	31.6	145	3 US-08-946-329A-69	Sequence 69, App1
23	44.5	31.6	297	3 US-09-058-489-10	Sequence 69, App1
24	44.5	31.6	376	4 US-09-849-016-8106	Sequence 8106, App
25	44.5	31.6	404	4 US-09-849-016-6202	Sequence 6202, App
26	44.5	31.6	416	4 US-09-949-016-6201	Sequence 6201, App
27	44.5	31.6	433	4 US-09-949-016-10900	Sequence 10900, A

28	44.5	31.6	471	1 US-08-176-427B-11	Sequence 11, App1
29	44.5	31.6	471	2 US-08-356-060A-34	Sequence 34, App1
30	44.5	31.6	471	3 US-08-460-900C-34	Sequence 34, App1
31	44.5	31.6	471	3 US-08-757-230A-8	Sequence 8, App1
32	44.5	31.6	471	3 US-08-674-509B-34	Sequence 34, App1
33	44.5	31.6	471	3 US-08-954-698-34	Sequence 34, App1
34	44.5	31.6	471	3 US-08-957-874-34	Sequence 34, App1
35	44.5	31.6	471	4 US-09-325-256-26	Sequence 26, App1
36	44.5	31.6	471	4 US-08-700-393-8	Sequence 8, App1
37	44.5	31.6	471	4 US-09-639-693-34	Sequence 34, App1
38	44.5	31.6	471	4 US-09-448-188-34	Sequence 34, App1
39	44.5	31.6	471	4 US-08-954-128-34	Sequence 34, App1
40	44.5	31.6	471	4 US-09-704-917-20	Sequence 20, App1
41	44.5	31.6	471	4 US-08-954-740-34	Sequence 34, App1
42	44.5	31.6	471	4 US-09-151-999-20	Sequence 20, App1
43	44.5	31.6	471	4 US-09-736-476-34	Sequence 34, App1
44	44.5	31.6	471	4 US-09-418-221-18	Sequence 18, App1
45	44.5	31.6	560	2 US-08-948-569A-10	Sequence 10, App1

ALIGNMENTS

```

RESULT 1
; Sequence 390, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 390
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-404-879A-390

Query Match          100.0%; Score 141; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db      334 DRVAIYEEFLRMTRNGTQLONFLLDRSS 361

RESULT 2
; Sequence 390, Application US/09667857
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455

```

```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 390
/ LENGTH: 438
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-667-857-390

Query Match          100.0%; Score 141; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 1,1e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIEEFLMTRNGTOLQNFLLDRSS 28
Db 334 DRVAIEEFLMTRNGTOLQNFLLDRSS 361

RESULT 3
US-09-404-879A-389
/ Sequence 389, Application US/09404879A
/ Patent No. 6468546
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C2
/ CURRENT APPLICATION NUMBER: US/09/404,879A
/ CURRENT FILING DATE: 1999-09-24
/ NUMBER OF SEQ ID NOS: 393
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 389
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-404-879A-389

Query Match          100.0%; Score 141; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 2,3e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIEEFLMTRNGTOLQNFLLDRSS 28
Db 729 DRVAIEEFLMTRNGTOLQNFLLDRSS 756

RESULT 4
US-09-667-857-389
/ Sequence 389, Application US/09667857
/ Patent No. 6699664
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C5
/ CURRENT APPLICATION NUMBER: US/09/667,857
/ CURRENT FILING DATE: 2000-09-20
/ NUMBER OF SEQ ID NOS: 455
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 389
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-667-857-389
```

```
Query Match          100.0%; Score 141; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 2,3e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIEEFLMTRNGTOLQNFLLDRSS 28
Db 729 DRVAIEEFLMTRNGTOLQNFLLDRSS 756

RESULT 5
US-09-404-879A-312
/ Sequence 312, Application US/09404879A
/ Patent No. 6468546
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C2
/ CURRENT APPLICATION NUMBER: US/09/404,879A
/ CURRENT FILING DATE: 1999-09-24
/ NUMBER OF SEQ ID NOS: 393
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-404-879A-312

Query Match          100.0%; Score 141; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2,6e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIEEFLMTRNGTOLQNFLLDRSS 28
Db 810 DRVAIEEFLMTRNGTOLQNFLLDRSS 837

RESULT 6
US-09-338-933-312
/ Sequence 312, Application US/09338933
/ Patent No. 648931
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer Lynn
/ APPLICANT: King, Gordon E.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
/ TITLE OF INVENTION: OVARIAN CANCER
/ FILE REFERENCE: 210121.462C1
/ CURRENT APPLICATION NUMBER: US/09/338,933
/ CURRENT FILING DATE: 1999-06-23
/ NUMBER OF SEQ ID NOS: 312
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-338-933-312

Query Match          100.0%; Score 141; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2,6e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIEEFLMTRNGTOLQNFLLDRSS 28
Db 810 DRVAIEEFLMTRNGTOLQNFLLDRSS 837

RESULT 7
US-09-667-857-312
/ Sequence 312, Application US/09667857
/ Patent No. 6699664
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Ranger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C5
/ CURRENT APPLICATION NUMBER: US/09/667,857
/ CURRENT FILING DATE: 2000-09-20
/ NUMBER OF SEQ ID NOS: 455
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapien
US-09-667-857-312
```

```
Query Match          100.0%; Score 141; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 28
    |||||:|||||:|||||:|||||:
DB 810 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 837
```

```
RESULT 8
US-09-404-879A-388
/ Sequence 388, Application US/09404879A
/ Patent No. 6468546
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C2
/ CURRENT APPLICATION NUMBER: US/09/404,879A
/ CURRENT FILING DATE: 1999-09-24
/ NUMBER OF SEQ ID NOS: 393
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 388
/ LENGTH: 772
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-404-879A-388
```

```
Query Match          39.7%; Score 56; DB 4; Length 772;
Best Local Similarity 42.9%; Pred. No. 1.6;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 28
    |||||:|||||:|||||:|||||:
DB 281 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 308
```

```
RESULT 9
US-09-667-857-388
/ Sequence 388, Application US/09667857
/ Patent No. 6699664
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Fling, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Ranger, Gary Richard
/ APPLICANT: Reed, Steven G.
```

```
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C5
/ CURRENT APPLICATION NUMBER: US/09/667,857
/ CURRENT FILING DATE: 2000-09-20
/ NUMBER OF SEQ ID NOS: 455
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 388
/ LENGTH: 772
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-667-857-388
```

```
Query Match          39.7%; Score 56; DB 4; Length 772;
Best Local Similarity 42.9%; Pred. No. 1.6;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 28
    |||||:|||||:|||||:|||||:
DB 281 DRAVAYEEFLRMTRNGTOLQNFITLDRSS 308
```

```
RESULT 10
US-09-710-279-332
/ Sequence 332, Application US/09710279
/ Patent No. 6703492
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/09/710,279
/ CURRENT FILING DATE: 2000-11-09
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 332
/ LENGTH: 263
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-332
```

```
Query Match          34.4%; Score 48.5; DB 4; Length 263;
Best Local Similarity 37.9%; Pred. No. 7.1;
Matches 11; Conservative 7; Mismatches 8; Indels 3; Gaps 1;
```

```
QY 3 VAIVYEEFLRMTRNGTOLQNFITLDRSS 28
    |||||:|||||:|||||:|||||:
DB 33 VAIFFKSPVLTTRDGEKTLQVALQRES 61
```

```
RESULT 11
US-09-134-001C-3921
/ Sequence 3921, Application US/09134001C
/ Patent No. 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14
/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NO 3921
```



```

/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD/ROM ISO9660
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Arinello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 5481:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 115 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (8) LOCATION 1...115
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5481:
US-09-107-532A-5481

Query Match      32.6%; Score 46; DB 4; Length 115;
Best Local Similarity 34.6%; Pred. No. 6.7;
Matches 9; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY      3 VAIYEEFLMTNRNGTOLONTLDPRS 28
DB      28 VAVYQAFKTPDYGTLLSLTHNQTA 53

```

Search completed: October 27, 2005, 17:34:35
 Job time : 25.3535 secs

COPIES OF THE

Result No.	Score	Query Match	Length	DB	ID	Description
1	141	100.0	28	15	US-10-612-090-18	Sequence 18, Appl
2	141	100.0	108	15	US-10-612-090-2	Sequence 2, Appl
3	141	100.0	108	15	US-10-612-090-13	Sequence 13, Appl
4	141	100.0	174	15	US-10-383-368-4	Sequence 4, Appl
5	141	100.0	178	9	US-09-684-441-489	Sequence 489, App
6	141	100.0	178	10	US-09-907-969-489	Sequence 489, App
7	141	100.0	178	14	US-10-196-053-489	Sequence 489, App
8	141	100.0	178	17	US-10-860-790-489	Sequence 489, App
9	141	100.0	233	9	US-09-684-441-488	Sequence 488, App
10	141	100.0	233	10	US-09-907-969-488	Sequence 488, App
11	141	100.0	233	14	US-10-196-053-488	Sequence 488, App

12	141	100.0	233	17	US-09-860-790-408	Sequence 488, App
13	141	100.0	284	10	US-09-965-738-300	Sequence 300, App
14	141	100.0	318	10	US-09-907-969-594	Sequence 594, App
15	141	100.0	318	14	US-10-198-053-594	Sequence 594, App
16	141	100.0	318	17	US-10-860-790-594	Sequence 594, App
17	141	100.0	367	15	US-10-333-900-27	Sequence 27, App
18	141	100.0	438	9	US-09-884-441-590	Sequence 390, App
19	141	100.0	438	9	US-09-884-441-483	Sequence 483, App
20	141	100.0	438	10	US-09-907-969-390	Sequence 390, App
21	141	100.0	438	10	US-09-907-969-483	Sequence 483, App
22	141	100.0	438	10	US-09-827-271-390	Sequence 390, App
23	141	100.0	438	14	US-10-198-053-390	Sequence 390, App
24	141	100.0	438	14	US-10-198-053-483	Sequence 483, App
25	141	100.0	438	15	US-10-333-900-19	Sequence 19, App
26	141	100.0	438	17	US-10-860-790-390	Sequence 390, App
27	141	100.0	438	17	US-10-860-790-483	Sequence 483, App
28	141	100.0	439	10	US-09-865-738-148	Sequence 148, App
29	141	100.0	526	15	US-10-333-900-30	Sequence 30, App
30	141	100.0	545	14	US-10-243-243-34	Sequence 4, App
31	141	100.0	583	14	US-10-142-515-4	Sequence 4, App
32	141	100.0	748	17	US-10-687-035-1	Sequence 1, App
33	141	100.0	809	17	US-10-687-035-2	Sequence 2, App
34	141	100.0	833	9	US-09-884-441-389	Sequence 389, App
35	141	100.0	833	10	US-09-907-969-389	Sequence 389, App
36	141	100.0	833	10	US-09-827-271-389	Sequence 389, App
37	141	100.0	833	14	US-10-198-053-389	Sequence 389, App
38	141	100.0	833	17	US-10-860-790-389	Sequence 389, App
39	141	100.0	914	9	US-09-778-320-206	Sequence 206, App
40	141	100.0	914	9	US-09-910-689-206	Sequence 206, App
41	141	100.0	914	9	US-09-884-441-512	Sequence 312, App
42	141	100.0	914	9	US-09-884-441-512	Sequence 478, App
43	141	100.0	914	10	US-09-907-969-312	Sequence 312, App
44	141	100.0	914	10	US-09-907-969-478	Sequence 478, App
45	141	100.0	914	10	US-09-827-271-312	Sequence 312, App

ALIGNMENTS

```

RESULT 1
US-10-612-090-18
: Sequence 18, Application US/10612090
: Publication NO. US20040057952A1
: GENERAL INFORMATION:
: APPLICANT: Immunogen, Inc.
: TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
: FILE REFERENCE: A8340
: CURRENT APPLICATION NUMBER: US/10/612,090
: CURRENT FILING DATE: 2003-07-03
: PRIOR APPLICATION NUMBER: US 60/393,094
: PRIOR FILING DATE: 2002-07-03
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 18
: LENGTH: 28
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-612-090-18

```

```

Query Match          100.0%; Score 141; DB 15; Length 28;
Best Local Similarity 100.0%; Prcd. No. 9, 2e-15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  DRAVAIYEEFLMTRNGTOLONFTLDRSS 28
        |||||
Db       1  DRAVAIYEEFLMTRNGTOLONFTLDRSS 28
        |||||

RESULT 2
US-10-612-090--2
; Sequence 2, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-090-2

Query Match          100.0%; Score 141; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 4,4e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEEFLMTRNGTOLQNFLLDRSS 28
    |||||
    60 DRVAIYEEFLMTRNGTOLQNFLLDRSS 87

RESULT 3
US-10-612-090-13
; Sequence 13, Application US/10612090
; Publication No. US20040057952A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)-(1)
; OTHER INFORMATION: Glutathione S-transferase fusion site
US-10-612-090-13

Query Match          100.0%; Score 141; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 4,4e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEEFLMTRNGTOLQNFLLDRSS 28
    |||||
    60 DRVAIYEEFLMTRNGTOLQNFLLDRSS 87

RESULT 4
US-10-383-368-4
; Sequence 4, Application US/10383368
; Publication No. US20040002091A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul E.
; APPLICANT: Ebner, Reinhard
; APPLICANT: Strovel, Jeffrey W.
; TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
; FILE REFERENCE: 689290-123
; CURRENT APPLICATION NUMBER: US/10/383,368
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US 60/362,527
; PRIOR FILING DATE: 2002-03-07
; ORGANISM: Homo sapiens
```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-383-368-4

Query Match          100.0%; Score 141; DB 15; Length 174;
Best Local Similarity 100.0%; Pred. No. 7,7e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEEFLMTRNGTOLQNFLLDRSS 28
    |||||
    70 DRVAIYEEFLMTRNGTOLQNFLLDRSS 97

RESULT 5
US-09-884-441-489
; Sequence 489, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-489

Query Match          100.0%; Score 141; DB 9; Length 178;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEEFLMTRNGTOLQNFLLDRSS 28
    |||||
    129 DRVAIYEEFLMTRNGTOLQNFLLDRSS 156

RESULT 6
US-09-907-969-489
; Sequence 489, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
```


US-09-907-969-489

Query Match 100.0%; Score 141; DB 10; Length 178;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 129 DRVAIYEEFLRMTRNGTQLONFLLDRSS 156

RESULT 7

US-10-198-053-489
; Sequence 489, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-489

Query Match 100.0%; Score 141; DB 14; Length 178;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 129 DRVAIYEEFLRMTRNGTQLONFLLDRSS 156

RESULT 8

US-10-860-790-489
; Sequence 489, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-790-489

Query Match 100.0%; Score 141; DB 17; Length 178;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 129 DRVAIYEEFLRMTRNGTQLONFLLDRSS 156

RESULT 9

US-09-884-441-488
; Sequence 488, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-488

Query Match 100.0%; Score 141; DB 9; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 129 DRVAIYEEFLRMTRNGTQLONFLLDRSS 156

RESULT 10

US-09-907-969-488
; Sequence 488, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-969-488

Query Match 100.0%; Score 141; DB 10; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVAIYEEFLRMTRNGTQLONFLLDRSS 28
Db 129 DRVAIYEEFLRMTRNGTQLONFLLDRSS 156

RESULT 11

US-10-198-053-488
; Sequence 488, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.

```

; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-488

Query Match      100.0%; Score 141; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 1,1e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DRAVAYEEFLMTRNGTOLQNFLLDRSS 28
Db      129 DRAVAYEEFLMTRNGTOLQNFLLDRSS 156

RESULT 12
US-10-860-790-488
; Sequence 488, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-790-488

Query Match      100.0%; Score 141; DB 17; Length 233;
Best Local Similarity 100.0%; Pred. No. 1,1e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DRAVAYEEFLMTRNGTOLQNFLLDRSS 28
Db      129 DRAVAYEEFLMTRNGTOLQNFLLDRSS 156

RESULT 13
US-09-965-738-300
; Sequence 300, Application US/09965738
; Publication No. US20030143667A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy
; TITLE OF INVENTION: Repeat Sequences of the CA125 Gene and Their Use for Diagnostic a
; FILE REFERENCE: 40715-258841
; CURRENT APPLICATION NUMBER: US/09/965,738
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/284,175
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 300
; LENGTH: 284

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-738-300

Query Match      100.0%; Score 141; DB 10; Length 284;
Best Local Similarity 100.0%; Pred. No. 1,4e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DRAVAYEEFLMTRNGTOLQNFLLDRSS 28
Db      180 DRAVAYEEFLMTRNGTOLQNFLLDRSS 207

RESULT 14
US-09-907-969-594
; Sequence 594, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-09-907-969-594

Query Match      100.0%; Score 141; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 1,6e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DRAVAYEEFLMTRNGTOLQNFLLDRSS 28
Db      214 DRAVAYEEFLMTRNGTOLQNFLLDRSS 241

RESULT 15
US-10-198-053-594
; Sequence 594, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 318

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: 136,248,268
; OTHER INFORMATION: Xaa = Any amino acid
US-10-198-053-594

```

```

Query Match      100.0%; Score 141; DB 14; Length 318;
Best local Similarity . 100.0%; Pred. No. 1,6e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 DRVAIYEEFLRMTRNGTOLQNFLLDRSS 28
      |||||
Db      214 DRVAIYEEFLRMTRNGTOLQNFLLDRSS 241

```

Search completed: October 27, 2005, 18:31:47
 Job time : 89.6 secs

The Negro Community

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: October 27, 2005, 18:13:49 ; Search time 17.3209 Seconds
(without alignments)
155.538 Million cell updates/sec

Title: US-10-612-090-18

Perfect score: 141
Sequence: 1 DRVAIYEEFLRMTRNGTQLQNFPLDRSS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53.5	37.9	311	2 AG1271	hypothetical prote
2	53.5	37.9	311	2 A11633	hypothetical prote
3	49	34.8	540	1 T01989	calium-dependent
4	48	34.0	281	2 AD2312	hypothetical prote
5	48	34.0	384	2 G82670	general secretory
6	48	34.0	576	2 A84902	auxin-regulated pr
7	47.5	33.7	466	1 B43332	glutamate decarbox
8	47.5	33.7	466	1 S24234	glutamate decarbox
9	47.5	33.7	466	1 B91178	glutamate decarbox
10	47.5	33.7	466	2 F86024	glutamate decarbox
11	47.5	33.7	466	2 G85726	glutamate decarbox
12	47.5	33.7	466	2 B90891	glutamate decarbox
13	47	33.3	231	2 T32146	hypothetical prote
14	47	33.3	316	2 AB2868	UDP-hexose transfe
15	47	33.3	316	2 F97644	UDP-hexose transfe
16	47	33.3	530	2 A86433	T518.19 protein -
17	47	33.3	2890	2 B71846	dna-directed RNA p
18	47	33.3	2890	2 F64669	DNA-directed RNA p
19	46.5	33.0	397	2 D96580	hypothetical prote
20	46	32.6	113	2 H86296	F309.1 protein - A
21	46	32.6	249	2 T35589	probable secreted
22	46	32.6	514	2 T10938	calium-dependent
23	46	32.6	725	1 Z1B222	gene 1 protein - p
24	45.5	32.3	862	2 E84507	hypothetical prote
25	45	31.9	71	2 G97985	hypothetical prote
26	45	31.9	272	2 T15351	hypothetical prote
27	45	31.9	275	2 S55978	hypothetical prote
28	45	31.9	286	2 T21545	hypothetical prote
29	45	31.9	314	2 C97332	ABC-type transpo

30	45	31.9	367	2 F96827	protein F20B17.8 l
31	45	31.9	391	2 T35698	hypothetical prote
32	45	31.9	516	2 JC6015	L- amino-acid oxid
33	45	31.9	538	2 C71946	CTP synthase - Hel
34	45	31.9	538	2 B64563	CTP synthase - H
35	45	31.9	591	2 S73708	MG31 homolog H08
36	44.5	31.6	105	2 F90939	hypothetical prote
37	44.5	31.6	105	2 B85788	hypothetical prote
38	44.5	31.6	297	1 D64724	carnitine racemase
39	44.5	31.6	297	2 G85484	carnitine racemase
40	44.5	31.6	297	2 G90633	carnitine racemase
41	44.5	31.6	471	2 A46400	segment polarity p
42	44	31.2	152	2 A10726	probable membrane
43	44	31.2	267	2 A11249	conserved hypotet
44	44	31.2	267	2 AB1612	conserved hypotet
45	44	31.2	330	2 G95187	conserved domain p

ALIGNMENTS

RESULT 1

AG1271
hypothetical protein lmo1575 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AG1271

R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihl, H.;

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Mat

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1271

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <Gla>

A:Cross-references: UNIPROT:Q8Y6V6; GB:NC_003210; PIDN:CAC99653.1; PID:g16411004; GSPDB:C

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1575

C:Superfamily: Mycoplasma conserved hypothetical protein MG190

Query Match 37.9%; Score 53.5; DB 2; Length 311;

Best Local Similarity 37.5%; Pred. No. 1.8;

Matches 12; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

QY 1 DRVAIYEEFLRMTRN-----GTQLQNFPLDRS 27
Db 181 DRPALYRELYPEKNTVKSGLITLQNFWDEN 212

RESULT 2

AG1633
hypothetical protein homolog l1m1610 [imported] - Listeria innocua (strain Ctip11262)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AG1633

R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihl, H.;

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Mat

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1633

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <Gla>

A:Cross-references: UNIPROT:Q92BE0; GB:AL592022; PIDN:CAC96841.1; PID:g16414097; GSPDB:G

Best Local Similarity 45.8%; Pred. No. 26;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 1 DVAIYEFRLRMTRNGTOLONFTL 24
DB 9 DNMRVIDEFDEMRNAHVOVKOTL 32

RESULT 7

B43332
glutamate decarboxylase (EC 4.1.1.15) beta - Escherichia coli (strain K-12)
N:Alternate names: L-glutamate 1-carboxy-lyase
C:Species: Escherichia coli
C:Date: 10-Mar-1994 #sequence_revision 23-Mar-1995 #ext_change 09-Jul-2004
C:Accession: B43332; S30261; H64902
R:Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A:Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to d
A:Reference number: A43332; MUID:92394884; PMID:1522060
A:Accession: B43332
A:Molecule type: DNA
A:Residues: 1-466 <SMI>

A:Cross-references: UNIPROT:P28302; GB:M84025; NID:G146059; PIDN:AAA23834.1; PID:G146060
R:Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A:Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A:Reference number: S30261; MUID:93204884; PMID:8455549
A:Accession: S30261

A:Molecule type: protein
A:Residues: 1-4, 'LOVL', 7-15 <YOS>
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64902
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-466 <BLAT>
A:Cross-references: GB:AE000246; GB:U00096; NID:G1787764; PIDN:AACT4566.1; PID:G1787769;
C:Genetics:
A:Gene: gadB
A:Map position: 78 min
C:Function:

A:Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A>Note: In E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
C:Superfamily: Escherichia coli glutamate decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal P
F/276/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 33.7%; Score 47.5; DB 1; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAOYEFRLRGREGITKVN 344

RESULT 8

S24234
glutamate decarboxylase (EC 4.1.1.15) alpha - Escherichia coli (strain K-12)
N:Alternate names: GAD alpha protein; L-glutamate 1-carboxy-lyase
C:Species: Escherichia coli
C:Date: 22-Jan-1993 #sequence_revision 05-Dec-1997 #ext_change 09-Jul-2004
C:Accession: S47737; H65149; S24234; S24241; A43332; PNO616
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47737
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-466 <PLU>
A:Cross-references: UNIPROT:P80063; EMBL:U00039; NID:G466582; PIDN:AA818493.1; PID:G4666
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65149
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-466 <BLAT>
A:Cross-references: GB:AE000428; GB:U00096; NID:G1789931; PIDN:AACT6542.1; PID:G1789934;
A:Experimental source: strain K-12, substrain MG1655
R:Riars, B.; Sweeney, G.; Barra, D.; Bossa, F.; John, R.A.
Eur. J. Biochem. 204, 93-98, 1992
A:Title: The amino acid sequence of glutamate decarboxylase from Escherichia coli. Evolut
A:Reference number: S24241; MUID:92155241; PMID:1740158
A:Accession: S24234

A:Molecule type: protein
A:Residues: 1-63, 'S', 65-72, 'R', 74-152, 'N', 154-164, 'S', 166-354, 'N', 356-466 <MAR>
A:Accession: S23421
A:Molecule type: DNA
A:Residues: 148-164, 'S', 166-207, 'N', 209-294, 'V', 296-431 <MAR1>
A:Cross-references: GB:X63123; NID:G41601; PIDN:CAA44834.1; PID:G938166
R:Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A:Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to d
A:Reference number: A43332; MUID:92394884; PMID:1522060
A:Accession: A43332

A:Molecule type: DNA
A:Residues: 1-466 <SMI>
A:Cross-references: GB:M84024; NID:G146057; PIDN:AAA23833.1; PID:G146058
A>Note: sequence extracted from NCBI backbone (NCBIN:112979; NCBI:P:112980)
R:Yoshida, T.; Yamashino, T.; Ueguchi, C.; Mizuno, T.
Biosci. Biotechnol. Biochem. 57, 1568-1569, 1993
A:Title: Expression of the Escherichia coli dimorphic glutamic acid decarboxylases is reg
A:Reference number: PNO616; MUID:94033862; PMID:7764225
A:Accession: PNO616
A:Molecule type: protein
A:Residues: 1-21, 'X' <YOS>
C:Genetics:
A:Gene: gadA
A:Map position: 78 min
C:Function:

A:Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A>Note: In E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
C:Superfamily: Escherichia coli glutamate decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal p
F/276/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 33.7%; Score 47.5; DB 1; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAOYEFRLRGREGITKVN 344

RESULT 9

E91178
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #ext_change 09-Jul-2004
C:Accession: E91178
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A89623; MUID:21156231; PMID:11258796
A:Accession: E91178
A:Status: preliminary
A:Molecule type: DNA

A:Reference number: A89623; MUID:21156231; PMID:11258796
A:Accession: E91178
A:Status: preliminary
A:Molecule type: DNA

```

A:Residues: 1-466 <HAV>
A:Cross-references: UNIPROT:P58228, GB:BA000007, PIDN:BAB37820.1, PID:G13363871, GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC64397
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          33.7%; Score 47.5; DB 2; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAQYEFRLRGEGYTKVON 344

RESULT 10
P86024
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: P86024
R:Perla, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: P86024
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <STO>
A:Cross-references: UNIPROT:P58228, GB:AE005174, NID:G12518229, PIDN:AAG58658.1, GSPDB:G
C:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gadA
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          33.7%; Score 47.5; DB 2; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAQYEFRLRGEGYTKVON 344

RESULT 11
G85726
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G85726
R:Perla, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: G85726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <STO>
A:Cross-references: UNIPROT:P28302, GB:AB005174, NID:G12515184, PIDN:AAG56275.1, GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gadB
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          33.7%; Score 47.5; DB 2; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAQYEFRLRMTRNG-TOLON 21

```

```

DB 325 IAQYEFRLRGEGYTKVON 344

RESULT 12
B90891
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90891
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Sasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hatori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <HAV>
A:Cross-references: UNIPROT:P28302, GB:BA000007, PIDN:BAB35521.1, PID:G13361564, GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC62098
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          33.7%; Score 47.5; DB 2; Length 466;
Best Local Similarity 55.0%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 VAIEEFLRMTRNG-TOLON 21
DB 325 IAQYEFRLRGEGYTKVON 344

RESULT 13
T32146
hypochemical protein Cl3A2.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32146
R:Rohlfing, T.; Wohlmann, P.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid Cl3A2.
A:Reference number: Z21126
A:Accession: T32146
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-231 <ROH>
A:Cross-references: UNIPROT:O16875, EMBL:AF022967, PIDN:AB69881.1, GSPDB:GN00023, CESP:
A:Experimental source: strain Bristol N2, clone Cl3A2
C:Genetics:
A:Gene: CESP:Cl3A2.11
A:Map position: 5
A:Introns: 7/3; 33/3; 71/2; 219/1

Query Match          33.3%; Score 47; DB 2; Length 231;
Best Local Similarity 47.4%; Pred. No. 13;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 9 FLBMTRNGTOLONFLDLS 27
DB 4 FVKRTONGKLKPNFVTRRS 22

RESULT 14
AB2868
UDP-hexose transferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2868
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavdin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

```


A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2868
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-316 <KUR>
A:Cross-references: UNIPROT:O8UCM1; GB:AE008688; PIDN:AAL43360.1; PID:g17740855; GSPDB:C
C:Genetics:
A:Gene: exu
A:Map position: circular chromosome
C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match 33.3%; Score 47; DB 2; Length 316;
Best Local Similarity 38.5%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 DRAIYEEFLRMTRNGTOLQNFLLDR 26
| | | : : : | : : : |
DB 123 DVVAVYGDYERIDENGTKIGRRNLIR 148

RESULT 15
F97644
UDP-hexose transferase (AF039306) [imported] - *Agrobacterium tumefaciens* (strain C58, C5
C:Species: *Agrobacterium tumefaciens*
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97644
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F97644
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-316 <KUR>
A:Cross-references: UNIPROT:O8UCM1; GB:AE007869; PIDN:AAK8811.1; PID:g15157543; GSPDB:C
C:Genetics:
A:Gene: AGR_C_4307
A:Map position: circular chromosome
C:Superfamily: Neisseria meningitidis glycosyl transferase A

Query Match 33.3%; Score 47; DB 2; Length 316;
Best Local Similarity 38.5%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 DRAIYEEFLRMTRNGTOLQNFLLDR 26
| | | : : : | : : : |
DB 123 DVVAVYGDYERIDENGTKIGRRNLIR 148

Search completed: October 27, 2005, 18:57:28
Job time : 18.3209 secs

1830-1831

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 17:34:49 ; Search time 81.3953 Seconds
(without alignments)
176.155 Million cell updates/sec

Title: US-10-612-090-18
Perfect score: 141
Sequence: 1 DRVAIYEEFLRMTRNGTOLQNFTRDRSS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	100.0	1148	Q9H7S7	Q9H7S7 homo sapien
2	141	100.0	6995	Q96RK2	Q96RK2 homo sapien
3	141	100.0	22152	Q8WXT7	Q8WXT7 homo sapien
4	125	88.7	258	Q9D1H1	Q9D1H1 mus musculu
5	56	39.7	867	Q6ZQW5	Q6ZQW5 homo sapien
6	54	38.3	419	Q6UANG	Q6UANG tetradon n
7	53.5	37.9	311	Q8Y6V6	Q8Y6V6 listeria m
8	53.5	37.9	311	Q92BEO	Q92BEO listeria in
9	53.5	37.9	311	Q71292	Q71292 listeria m
10	53	37.6	497	Q726P9	Q726P9 desulfovibr
11	52	36.9	530	Q7Q949	Q7Q949 anopheles g
12	51.5	36.5	631	Q8SVW5	Q8SVW5 ancephalito
13	51	36.2	186	Q6SDU9	Q6SDU9 bacillus li
14	51	36.2	339	Q9FHS7	Q9FHS7 arabidopsis
15	51	36.2	591	Q9LWC3	Q9LWC3 oryza sativ
16	51	36.2	1059	Q7RKI8	Q7RKI8 plasmidium
17	50	35.5	388	Q74GS7	Q74GS7 geobacter s
18	49.5	35.1	870	Q7QAH7	Q7QAH7 anopheles g
19	49	34.8	540	Q81390	Q81390 nicotiana t
20	49	34.8	553	Q8QFT1	Q8QFT1 mugil cepha
21	49	34.8	766	POOP_PSEPK	POOP_PSEPK
22	48.5	34.4	273	Q8CKM1	Q8CKM1 pseudomonas
23	48	34.0	145	Q8PLV2	Q8PLV2 staphylococ
24	48	34.0	191	Q81JQ8	Q81JQ8 xanthomonas
25	48	34.0	248	Q9CQJ0	Q9CQJ0 m mus muscu
26	48	34.0	281	Q8YYP9	Q8YYP9 anabaena sp
27	48	34.0	282	Q76M63	Q76M63 pseudomonas
28	48	34.0	384	Q87DE6	Q87DE6 xylella fas
29	48	34.0	384	Q9PD55	Q9PD55 xylella fas
30	48	34.0	575	Q949V9	Q949V9 arabidopsis
31	48	34.0	575	Q9SKR2	Q9SKR2 arabidopsis

32	48	34.0	1109	Q6FJ53	Q6FJ53 candida gla
33	48	34.0	1789	Q7SGG0	Q7SGG0 neospora
34	47.5	33.7	466	DCEA_ECO57	P58228 escherichia
35	47.5	33.7	466	DCEA_ECOL1	P80063 escherichia
36	47.5	33.7	466	DCEB_ECOL1	P28302 escherichia
37	47.5	33.7	466	Q7UAY1	Q7UAY1 shigella fl
38	47.5	33.7	487	Q83PR1	Q83PR1 shigella fl
39	47.5	33.7	489	Q8FHS5	Q8FHS5 escherichia
40	47.5	33.7	572	Q9VXL9	Q9VXL9 drosophila
41	47.5	33.7	867	Q6L881	Q6L881 eutrepria v
42	47	33.3	179	Q6Z4B4	Q6Z4B4 oryza sativ
43	47	33.3	271	Q99JY9	Q99JY9 mus musculu
44	47	33.3	286	Q93TW5	Q93TW5 stigmella
45	47	33.3	316	Q8UCW1	Q8UCW1 agrobacteri

ALIGNMENTS

RESULT 1

Q9H7S7 PRELIMINARY; PRT; 1148 AA.

AC Q9H7S7, 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; DOI=10.1038/ng1285;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori S.,
RA Tani H., Kimata S., Watanabe S., Yoshida M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hiraoka S., Chiba Y.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Toguya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamaguchi T., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togoashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okunura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Iegai T., Sugano S.;
RT Complete sequencing and characterization of 21,243 full-length human
cDNAs.
RT Nat. Genet. 36:40-45(2004).
DR EMBL; AK024365; BAB14899.1; --
DR HSSP; O9DHL1; 11VZ.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 7.
DR PROSITE; PS50024; SEA; 3.
SQ SEQUENCE 1148 AA; 127957 MW; 3861BD5EFDFA8C CRC64;

```

Query Match          100.0%; Score 141; DB 2; Length 1148;
Best Local Similarity 100.0%; Pred. No. 7.6e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEFLLMTNNGTOLQNFLLDRSS 28
Db 1044 DRVAIYEFLLMTNNGTOLQNFLLDRSS 1071

RESULT 2
ID Q96RK2 PRELIMINARY; PRT; 6995 AA.
AC Q96RK2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mucin 16 (Fragment).
GN Name=MUC16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200;
RA Yin B.W., Lloyd K.O.;
RT "Molecular cloning of the ca125 ovarian cancer antigen. identification
as a new mucin, muc16."
RL J. Biol. Chem. 276:27371-27375(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Lloyd K.O., Yin B.W.T.;
RA Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases.
RL EMBL; AF161486; AAK74120.3; -.
DR HSSP; Q9DIH1; 11VZ.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 20.
DR SMART; SMO0200; SEA; 10.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50024; SEA; 6.
FT NON TER 1
SQ SEQUENCE 6995 AA; 744958 MM; 80C797DBDF33A2B CRC64;

Query Match          100.0%; Score 141; DB 2; Length 6995;
Best Local Similarity 100.0%; Pred. No. 6.1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEFLLMTNNGTOLQNFLLDRSS 28
Db 6891 DRVAIYEFLLMTNNGTOLQNFLLDRSS 6918

RESULT 3
ID Q8WX17 PRELIMINARY; PRT; 22152 AA.
AC Q8WX17;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Ovarian cancer related tumor marker CA125.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
RA O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RA York L.;
RT "The CA 125 gene: an extracellular superstructure dominated by repeat
sequences";
RL Tumour Biol. 22:348-366(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA O'Brien T.J., Underwood L.J., Beard J.B.;
RL Submitted (OCT-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF14442; AA16513.2; -.
DR Genew; HGNC:15582; MUC16.
DR Pfam; PF01390; SEA; 51.
DR SMART; SMO0200; SEA; 23.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_2.
DR PROSITE; PS50024; SEA; 11.
SQ SEQUENCE 22152 AA; 2352668 MM; B3E7BDF19997A440 CRC64;

Query Match          100.0%; Score 141; DB 2; Length 22152;
Best Local Similarity 100.0%; Pred. No. 2.3e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRVAIYEFLLMTNNGTOLQNFLLDRSS 28
Db 22048 DRVAIYEFLLMTNNGTOLQNFLLDRSS 22075

RESULT 4
ID Q9DIH1 PRELIMINARY; PRT; 258 AA.
AC Q9DIH1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA. RIKEN full-length enriched
library, clone:111000814 product:hypothetical SBA domain containing
protein, full insert sequence.
GN Name=111000814R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

```

```
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kanai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1157-1771(2000).
[6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi J., Aizawa K., Akahira S., Akiyama T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Furumishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imocani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawaji Y., Kojima Y., Komno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RU EMBL; AK003577; BAB22869.1; -.
DR PDB; 1IVZ; NMR; A=60-190.
DR MGD; MGI:1920982; 1110008114R1k.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 1.
DR PROSITE; PS50024; SEA; 1.
KM Hypothetical protein.
SQ SEQUENCE 258 AA; 29425 MW; B64D9B63394D84E7 CRC64;

Query Match
Best Local Similarity 89.7%; Score 125; DB 2; Length 258;
Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DRAVYEEFLMTRNGTOLQNFLLDRSS 28
Db 154 DRAVYEEFLMTRNGTOLQNFLLDRKS 181

RESULT 5
Q6ZOW5 PRELIMINARY; PRT; 867 AA.
AC Q6ZOW5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46845.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RP TISSUE=uterus;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohara N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Matenabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamauchi H., Matsuo K., Nakamura Y., Sakine M.,
RA Kikuchi H., Kanda K., Nagatsuna M., Murakawa K., Kanehori K.,
RA Takahashi-Fuji A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuko Y., Nagai K., Isegai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK126681; BAC87568.1; -.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 4.
DR PROSITE; PS50024; SEA; 2.
SQ SEQUENCE 867 AA; 96200 MW; C7B3033258839622 CRC64;
```

```
QY Query Match
Best Local Similarity 39.7%; Score 56; DB 2; Length 867;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 DRAVYEEFLMTRNGTOLQNFLLDRSS 28
Db 380 DRAVYEEFLMTRNGTOLQNFLLDRSS 407

RESULT 6
Q6UN6 PRELIMINARY; PRT; 419 AA.
AC Q6UN6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Class I helical cytokine receptor number 20.
GN Name=CRPA20;
OS Tetradon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxID=99883;
RN (1)
RP SEQUENCE FROM N.A.
RA Tailon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthonard V., Jubin C., Castelli V., Karkina M., Vacherie B.,
RA Blumont C., Skalli Z., Catroico L., Poulin J., Berardins Vd.,
RA Craud C., Duprat S., Brotier P., Coutanceau J.-P., Couzy J.,
RA Parra G., Lardier G., Chappe C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigs R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quittier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Crolius H.R.;
RT "Analysis of the Tetraodon nigroviridis genome reveals the
RT proteome of bony vertebrates and its duplication in teleost
RT fish."
RL Nature 0:0-0(2004).
DR EMBL; AY374492; AAR25683.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003531; Hemiptericeps_F1.
DR PROSITE; PS01355; HEMATOPOI_REC_S_F1; UNKNOWN_1.
KM Receptor.
SQ SEQUENCE 419 AA; 46687 MW; 0F100D2F20EA4FF7 CRC64;

Query Match
Best Local Similarity 38.3%; Score 54; DB 2; Length 419;
Matches 14; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 6 YEEFLR---MTRNGT-OLQNFLLDRSS 28
Db 146 YEEFLR---MTRNGT-OLQNFLLDRSS 172

RESULT 7
Q8Y6V6 PRELIMINARY; PRT; 311 AA.
AC Q8Y6V6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE lmo1575 protein.
GN Ordered locus names=lmo1575;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxID=1639;
RN (1)
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21531279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
BA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Checouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud B., Durant L., Dussurget O.,
RA Entian K.-D., Felhi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunat F., Kurapkak G.,
RA Madueno E., Maicouram A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Pucell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RL "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
EMBL: AL591979; CAC99653.1; -.
DR PIR: A01271; A01271.
DR ListList: LM01575; -.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR Pfam: PF01368; DHH; 1.
DR Pfam: PF02272; DHHA1; 1.
KW Complete proteome.
SQ SEQUENCE 311 AA; 34799 MW; 2A38F91748618C8E CRC64;

Query Match 37.9%; Score 53.5; DB 2; Length 311;
Best Local Similarity 37.5%; Pred. No. 8.7;
Matches 12; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

OY 1 DRVAIYEEFLMTRN-----GTOLNFTLDRS 27
Db 181 DRPALRYELVELPKNTVKTLSGYILQNFWDEN 212

RESULT 8
O92BE0 PRELIMINARY; PRT; 311 AA.
ID O92BE0;
DC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ltn1610 protein.
GN OrderedLocustNames=ltn1610;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CX NCBI_TaxID=1642;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
BA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Checouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud B., Durant L., Dussurget O.,
RA Entian K.-D., Felhi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunat F., Kurapkak G.,
RA Madueno E., Maicouram A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Pucell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RL "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
EMBL: AL596169; CAC96841.1; -.
DR PIR: A11633; A11633.
DR ListList: Ltn1610; -.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR InterPro: IPR003156; Peptidase_DHHA1.
DR InterPro: IPR001667; Peptidase_Recl.
DR Pfam: PF01368; DHH; 1.
DR Pfam: PF02272; DHHA1; 1.
KW Complete proteome.
SQ SEQUENCE 311 AA; 34785 MW; 9C26D75EE426AE13 CRC64;

```

```

Query Match 37.9%; Score 53.5; DB 2; Length 311;
Best Local Similarity 37.5%; Pred. No. 8.7;
Matches 12; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

OY 1 DRVAIYEEFLMTRN-----GTOLNFTLDRS 27
Db 181 DRPALRYELVELPKNTVKTLSGYILQNFWDEN 212

RESULT 9
O71Z92 PRELIMINARY; PRT; 311 AA.
ID O71Z92;
DC 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE DHH subfamily 1 protein.
GN OrderedLocustNames=LM012365_1597;
OS Listeria monocytogenes (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
CX NCBI_TaxID=265669;
[1]
RP SEQUENCE FROM N.A.
RC PubMed=15115801; DOI=10.1093/nar/gkh562;
RX Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Raeko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nieman W.C., Beaman M.J.,
RA Brink L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Pedocova N.,
RA Forbester H.A., Tran B., Katharou S.M.,
RA Bayles D.O., Luchaneky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
borne pathogen Listeria monocytogenes reveal new insights into the
core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL: AE017337; AAT04372.1; -.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR InterPro: IPR003156; Peptidase_DHHA1.
DR InterPro: IPR001667; Peptidase_Recl.
DR Pfam: PF01368; DHH; 1.
DR Pfam: PF02272; DHHA1; 1.
KW Complete proteome.
SQ SEQUENCE 311 AA; 34928 MW; 78EF15BC7570B26B CRC64;

Query Match 37.9%; Score 53.5; DB 2; Length 311;
Best Local Similarity 37.5%; Pred. No. 8.7;
Matches 12; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

OY 1 DRVAIYEEFLMTRN-----GTOLNFTLDRS 27
Db 181 DRPALRYELVELPKNTVKTLSGYILQNFWDEN 212

RESULT 10
O726P9 PRELIMINARY; PRT; 497 AA.
ID O726P9;
DC 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase, putative.
GN OrderedLocustNames=DVU1057;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
8303)
OC Bacteria; Proteobacteria; Delta proteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
CX NCBI_TaxID=882;
[1]
RP SEQUENCE FROM N.A.
RC PubMed=15077118; DOI=10.1038/nbt959;
RX Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,

```

RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radue D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.,
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough".
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL: AE017319; AAC97528.1; -.
DR TIGR: DVU3057; -.
DR GO: GO:0005737; Cytoplasm; IEA.
DR GO: GO:0003824; Catalytic activity; IEA.
DR GO: GO:0004109; F:coproporphyrinogen oxidase activity; IEA.
DR GO: GO:0005506; P:iron ion binding; IEA.
DR GO: GO:0006779; P:porphyrin biosynthesis; IEA.
DR InterPro: IPR006638; ELP3/Mab/NiB.
DR InterPro: IPR010723; HemN_C.
DR InterPro: IPR004559; HemN_Tel.
DR InterPro: IPR000169; Pept_cys_acsite.
DR InterPro: IPR007197; Radical SAM.
DR InterPro: IPR011060; Rbp_bind_barrel.
DR Pfam: PF06969; HemN_C; 1.
DR Pfam: PF04055; Radical_SAM; 1.
DR SMART: SM00729; ELP3; 1.
DR TIGRFAMs: TIGR00539; hemN_tel; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 497 AA; 54126 MW; 11D323B1AFFA70DB CRC64;

Query Match 37.6%; Score 53; DB 2; Length 497;
Best Local Similarity 45.8%; Pred. No. 18;
Matches 11; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 VAAYEEFLMTNGTOLQNTLDR 26
Db 443 VNRIDGYLRKTRNGMLVNSILR 466

RESULT 11
Q70949 PRELIMINARY; PRT; 530 AA.
AC Q70949;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCC4398 (Fragment).
GN Name=agCG50325; ORFNames=ENSANG00000010154;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008905; EAA09676.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR007928; Antifreeze_CF.
DR InterPro: IPR002159; CD36.
DR Pfam: PF01130; CD36; 1.
DR PRINTS: PR01609; CD36FAMILY.
FT NON_TER 1
SQ SEQUENCE 530 AA; 59511 MW; 9F8C26FC42488FF7 CRC64;

Query Match 36.9%; Score 52; DB 2; Length 530;
Best Local Similarity 69.2%; Pred. No. 28;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 12 MTRNGTOLQNTL 24
Db 257 MTRNGTSLNENFTI 269

RESULT 12
Q85VW5 PRELIMINARY; PRT; 631 AA.
ID Q85VW5
AC Q85VW5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein EC004_0440.
GN Name=EC004_0440;
OS Encephalitozoon cuniculi GB-M1.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=284813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
RA Klinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prestier G., Barde V., Peyretallade E., Broctier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
Encephalitozoon cuniculi".
RT Nature 414:450-453(2001).
RL Nature 414:450-453(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA GenomeScope;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL590444; CAD25231.1; -.
KW Hypothetical protein.
SQ SEQUENCE 631 AA; 73302 MW; F8EB93FABE2C5F25 CRC64;

Query Match 36.5%; Score 51.5; DB 2; Length 631;
Best Local Similarity 39.3%; Pred. No. 41;
Matches 11; Conservative 6; Mismatches 4; Indels 7; Gaps 1;

Qy 1 DRAAYEEFLMTNGTOLQNTLDRSS 28
Db 375 DSIFVYKQFLNVI-----NFTLRRSS 395

RESULT 13
Q65DU9 PRELIMINARY; PRT; 186 AA.
ID Q65DU9
AC Q65DU9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YWKC (Putative DNA binding).
GN Name=ywkc; Synonyms=raca; ORFNames=BL04032, BL103951;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Metkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential".
RT J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,

